

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2003, 06:24:06 ; Search time 4227 Seconds
(without alignments)
4461.637 Million cell updates/sec

Title: US-09-043-944-1
Perfect score: 2378
Sequence: 1 MFSTRQEGGGADAEHTTV.....YYINSLFLPFLCIINFISIIS 461

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xl
-Q=/cgn2.1/USPTO.spool.p/US09043944/runat_11122003_062400_3671/app.query.fasta_1.647
-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=act -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcg -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09043944 @CGN_1.1.4958 @runat_11122003_062400_3671 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

RESULT 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	2362	99.3	1461	3	CEU35660
2	2263	95.2	1335	3	AF171064
3	2052.5	86.3	27102	3	U41540
4	1009.5	42.5	1836	5	AY043492
5	997	41.9	1896	4	AF038936
6	993.5	41.8	1488	6	AF086330
7	993.5	41.8	2764	6	AR095621
8	993.5	41.8	2765	6	AX244146
9	993.5	41.8	2765	6	AX244648
10	993.5	41.8	2765	9	HUMS182R
11	993.5	41.8	2765	11	G27112
12	993.5	41.8	2791	6	AR060154
13	993.5	41.8	2791	6	AR087309
14	993.5	41.8	2791	6	AR134496
15	993.5	41.8	2791	6	AR256768
16	993.5	41.8	2792	6	AR144597
17	993.5	41.8	2811	9	BC011729
18	992	41.7	1764	3	DMU78084
19	989.5	41.6	2791	6	AR060024
20	989.5	41.6	2791	6	AR087179
21	989.5	41.6	2791	6	AR134366
22	989.5	41.6	2791	6	AR144467
23	989.5	41.6	2791	6	AR256638
24	989.5	41.6	3056	9	HUMPS1MRNA
25	989.5	41.6	3087	6	AR060026
26	989.5	41.6	3087	6	AR087181
27	989.5	41.6	3087	6	AR134368
28	989.5	41.6	3087	6	AR144469
29	989.5	41.6	3087	6	AR256640
30	989	41.6	1962	6	AR060155
31	989	41.6	1962	6	AR087310
32	989	41.6	1962	6	AR134497
33	989	41.6	1962	6	AR256769
34	989	41.6	1964	6	AR144598
35	989	41.6	1964	10	MUS51PR
36	989	41.6	2016	10	BC030409
37	989	41.6	2681	6	AG3557
38	989	41.6	2681	6	AR271266
39	988.5	41.6	1809	3	DMU77934
40	988.5	41.6	1894	3	AF017024
41	988.5	41.6	1895	6	AR060183
42	988.5	41.6	1895	6	AR087338
43	988.5	41.6	1895	6	AR134525
44	988.5	41.6	1895	6	AR256797
45	988.5	41.6	1914	6	AR082573

ALIGNMENTS

CEU35660	1461 bp	mRNA	linear	INV 23-JAN-1996
LOCUS	Caenorhabditis elegans membrane protein (sel-12)	mrna	linear	INV 23-JAN-1996
DEFINITION	cds.			
ACCESSION	U35660			
VERSION	U35660.1	GI:1109793		
KEYWORDS	Caenorhabditis elegans			
SOURCE	Caenorhabditis elegans			
ORGANISM	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.			
REFERENCE	1 (bases 1 to 1461)			
AUTHORS	Leviton, D. and Greenwald, I.			
TITLE	Facilitation of lin-12-mediated signalling by sel-12, a Caenorhabditis elegans S182 Alzheimer's disease gene			
JOURNAL	Nature 377 (6547), 351-354 (1995)			
MEDLINE	96032531			
PUBMED	7566091			
REFERENCE	2 (bases 1 to 1461)			
AUTHORS	Leviton, D.			
TITLE	Direct Submission			
JOURNAL	Submitted (07-SEP-1995) Diane Levitan, Biochemistry, Columbia University, 630 W. 168th St., New York, NY 10032, USA			
COMMENT	On Dec 9, 1995 this sequence version replaced gi:1000566.			
FEATURES	Location/Qualifiers			
source	1..1461			
	/organism="Caenorhabditis elegans"			
	/mol_type="mRNA"			
	/strain="Bristol var N2"			
	/db_xref="taxon:6239"			
	/map="IGX"			
gene	1..1461			
	/genes="sel-12"			
CDS	1..1386			
	/genes="sel-12"			
	/note="multiple membrane spanning domains"			
	/codon_start=1			
	/product="membrane protein"			
	/protein_id="AAA85511.1"			
	/db_xref="GI:1109794"			
	/translation="MPSRRQEGGGADAETHVYTNLTITNRNSQEDENVVEAEFLK YGASHVILFVPSVLCMALVFTMTNTITFYQNNRGHLLYTPVRETDSIVKGLMSL GNALVHLVCLVMTLLVIFYKVKYKLLHGLWLVSSFLLEFLTIYVQVBLKSFV SPALLVFLGLNYGVLMCHHWKGLRLQFYLITMSALMALVFIKYLPEWTVNFV LFVSSVDVLVTPKGRURYLVETAQERNEPIFPALIIYSSGVIPYVLVTAVENTTD PRETSDDNTSTAFPEASCSETPRPKVKRIPOKQVIESNTTASITQNSGVVER ELAAERTVQDANFHEHEERGVKLGDFIFYSVLKQASSYFDWNTTIACYVAIL IGLCFTLVLLAVFKRALPALQFPSPDSFTTFVPAGSSPHLLHKSLSKSYVYINSFLP FLCIINFSIIS"			
polyA_site	1461			
	/genes="sel-12"			
	/note="20 A residues"			
BASE COUNT	379 a 309 c 310 g 463 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	2,19e-199	Length:	1461	
Score:	2362.00	Matches:	459	
Percent Similarity:	99.57%	Conservative:	0	
Best Local Similarity:	99.57%	Mismatches:	2	
Query Match:	99.33%	Indels:	0	
DB:	3	Gaps:	0	
US-09-043-944-1 (1-461) x CEU35660 (1-1461)				
QY	1	MetProSerThrArgArgGlnGlnGluGlyGlyValAlaAspAlaGluThrHisThrVal	20	
Db	1	ATGCGCTCCACAGAGGACACACAGAGGCGGAGGTGCAGATCGCGAAACATACCGTT	60	
QY	21	TyrGlyThrAsnLeuLeuThrAsnArgAsnSerGlnGluAspGluAsnValValGluGlu	40	
Db	61	TACGGGTACAAATCTGATAAAATCGAATAGCATAGCAGAGAGAGAGAGAGAGAGAG	120	
QY	401	LeuLeuAlaValPheLysArgAlaLeuProAlaLeuGlnPheProPheSerProAspSer	420	
QY	361	GlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSerSerTyrPheAsp	380	
Db	1081	GGTCTGGGCGACTTCATTTTCTACTCTGTTCTCTCTCGCAGAGGCTTCATCGTACTTTG	1140	
QY	381	TyrAsnThrThrIleAlaCysTyrValAlaIleLeuIleGlyLeuCysPheThrLeuVal	400	
Db	1141	TGGAACACAGCATATCGCTTGTTATGTGGCCATCTTATTCGGGTCTCTGCTTACTTTG	1200	
QY	401	LeuLeuAlaValPheLysArgAlaLeuProAlaLeuGlnPheProPheSerProAspSer	420	
QY	41	AlaGluLeuLysTyrGlyAlaSerHisValIleHisLeuPheValProValSerLeuCys	60	
Db	121	GCGGAGCTGAAATACGGAGCATCTCAGTTATTCACTATTGTGCGCGGTGCTACTATGC	180	
QY	61	MetAlaLeuValValPheThrMetAsnThrIleThrPheTyrSerGlnAsnAsnGlyArg	80	
Db	181	ATGGCTCTGGTTGTTTATACGATGAACACACGATTACGTTTATAGTCAAAACAATGGA	240	
QY	81	HisLeuLeuSerHisProPheValArgGluThrAspSerIleValGluLysGlyLeuMet	100	
Db	241	CAITTTACTATACACTCTCTTTGTCGGGAAACAGACAGTATCGTTGAGAGGGAATCAT	300	
QY	101	SerLeuGlyAsnAlaLeuValMetLeuCysValValValLeuMetThrValLeuLeuIle	120	
Db	301	TCACTTTGGAAATGCTCTCGTCATGTTGTGGTGGTCTGCTGATGACAGTTCTGCTGAT	360	
QY	121	ValPheTyrLysTyrLysPheTyrLysLeuIleHisGlyTyrLeuIleValSerSerPhe	140	
Db	361	GTTTTCTATAAATACAAAGTTTATAAGCTTATTATCATGGATGGCTTATTGTGACAG	420	
QY	141	LeuLeuLeuPheLeuPheThrThrIleTyrValGlnGluValLeuLysSerPheAspVal	160	
Db	421	CTTCTCTTTTCTCTATCTACTACATCTATGTCAGAGGTTCTGAAAGTTTTCGATGTG	480	
QY	161	SerProSerAlaLeuLeuValLeuPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMet	180	
Db	481	TCCTCCAGCGCACTATTGTTTGTGTTGTTGACATGAGTATCGAGTTCTCGGAATCATG	540	
QY	181	CysIleHisTyrLysGlyProLeuArgLeuGlnGlnPheTyrLeuIleThrMetSerAla	200	
Db	541	TGTATACATTGAAAGGTCATTTGCTGCTGCAACAGTTCTACCTTATTACAAATGCTGCA	600	
QY	201	LeuMetAlaLeuValPheIleLysTyrLeuProGluTyrThrValTrpPheValLeuPhe	220	
Db	601	CTAATGCTCTGCTGCTTTTATCAAGTACCTACCAAGATGAGCTGTGGTGTGTGCTGTT	660	
QY	221	ValIleSerValTrpAspLeuValAlaValLeuThrProLysGlyProLeuArgTyrLeu	240	
Db	661	GTATCTCGGTTTGGATCTGGTTGCGGTGCTCACACCAAGAGGACCATGAGATATTG	720	
QY	241	ValGluThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSerSerGly	260	
Db	721	GTGGAAATCGCACAGAGAGAAACAGGCAATTTTCCCGCGCTGATTATTTCGTCTGGA	780	
QY	261	ValIleTyrProTyrValLeuValThrAlaValGluAsnThrThrAspProArgGluPro	280	
Db	781	GTCATCTATCCCTACGTTCTTGTCTGAGTTGAAACACGACACCCCGGTGAACCG	840	
QY	281	ThrSerSerAspSerAsnThrSerThrAlaPheProGlyGluAlaSerCysSerSerGlu	300	
Db	841	ACGTGCTCAGACTCAAAATACTTCTACAGCTTTTCTGGAGAGCGGAGTTGTTTCATCTGA	900	
QY	301	ThrProLysArgProLysValLysArgIleProGlnLysValGlnIleGluSerAsnThr	320	
Db	901	ACGCCAAAACGGCCAAAAGTGAACGAATTCCTCAAAAAGTGCATAATCGAATCATCT	960	
QY	321	ThrAlaSerThrThrGlnAsnSerGlyValArgValGluArgGluLeuAlaGluArg	340	
Db	961	ACAGCTTCACACACACACAAAACCTCTGAGTPAAGGTGGAGCGGAGCTACTCTGAGAGA	1020	
QY	341	ProThrValGlnAspAlaAsnPheHisArgHisGluGluGluArgGlyValLysLeu	360	
Db	1021	CCAACTGTACAAGACGCCAAATTTTACAGGACGACGAGAGAGAGAGAGAGAGAGAG	1080	
QY	361	GlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSerSerTyrPheAsp	380	
Db	1081	GGTCTGGGCGACTTCATTTTCTACTCTGTTCTCTCTCGCAGAGGCTTCATCGTACTTTG	1140	
QY	381	TyrAsnThrThrIleAlaCysTyrValAlaIleLeuIleGlyLeuCysPheThrLeuVal	400	
Db	1141	TGGAACACAGCATATCGCTTGTTATGTGGCCATCTTATTCGGGTCTCTGCTTACTTTG	1200	
QY	401	LeuLeuAlaValPheLysArgAlaLeuProAlaLeuGlnPheProPheSerProAspSer	420	

```

|||||
1201 CTGCTCGCGTCTTCAACAGCACTCCCGGTCTGCAATTTCCATTTCTCGGACTCA 1260
Qy 421 PhePheThrPheValProAlaGlySerSerProHisLeuLeuHisGlySerLeuLysSer 440
Db 1261 TTTTCTACTTTTGTACCGCTGGATCATCACCCCATTTCTTACACAAGTCTCTCAAAAGT 1320
Qy 441 ValTyTyTyrIleAsnSerLeuPheLeuProPheLeuCysIleIleAsnPheSerIleIle 460
Db 1321 GTTTATTATTAATTTCTCTGTTTTGTGCAATTTCTTGCATCATCAACTTTTCGATTATA 1380
Qy 461 Ser 461
Db 1381 TCT 1383

RESULT 2
AF171064 1335 bp mRNA linear INV 25-OCT-2002
LOCUS Caenorhabditis elegans presentin SEL-12 (sel-12) mRNA, complete
DEFINITION cds.
ACCESSION AF171064 GI:5759132
VERSION AF171064.1
KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida;
Rhabditoidea; Rhabditiidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 1335)
Wittenburg,N., Eimer,S., Lakowski,B., Rohrig,S., Rudolph,C. and
Baumeister,R.
Presentin is required for proper morphology and function of
neurons in C. elegans
NATURE 406 (6793), 306-309 (2000)
20372200
PUBMED 10917532
REFERENCE
2 (bases 1 to 1335)
Eimer,S. and Baumeister,R.
Direct Submission
TITLES
Submitted (22-JUN-1999) Genzentrum, LMU Munich, Feodor-Lynen-Str.
25, Munich 81377, Germany
FEATURES
source
1..1335
Location/Qualifiers
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/chromosome="X"
1..1335
/gene="sel-12"
1..1335
/gene="sel-12"
/function="facilitator of Notch receptors signaling"
/note="membrane protein; similar to Homo sapiens PS1 and
PS2; similar to SEL12 deposited in GenBank Accession
Number AAA85511"
/codon_start=1
/product="presentin SEL-12"
/protein_id="AAD50991.1"
/db_xref="GI:5759133"
/translation="MPSTRQEGGGDAETHVTGNTLITNRNSQEDENVVBAELK
YGASHVILFVPSLCLMALVFTWNTITFYSONNGRHLLYTPFVRETDTSIVKGLMSL
GNALVLCVVLMVTLVLYFYKPKYKLIHGLWLVSSFLLLFLFTTIYQVBLKSFV
SPSALLVFLGNGVGLMNCIHWKGPLRLQQLITNSALMALVFIKILPEWTFV
LFVLSMDLVAVLTPKGPLRYLVETAQERNEPIFPALYSSGVYPPYLVTAVENTTD
PREPSSDSTNFPASCSSETPKPKVRIPOKQVIESNTASTTQNSGVYRVER
ELAAERPTVQDANFHRHEERGVKGLGDFIFYSLGKASSYFDWNTTIACYVAIL
IGLCFTLVLLAVFKRALPALPISIFPSGLIFEFCTRWIITPFTVQVSQCLLY"
BASE COUNT 345 a 291 c 301 g 398 t
ORIGIN
Alignment Scores:
Pred. No.: 1.11e-190 Length: 1335
Score: 2263.00 Matches: 442

```

```

Percent Similarity: 99.33% Conservative: 0
Best Local Similarity: 99.33% Mismatches: 2
Query Match: 95.16% Indels: 1
DB: 3 Gaps: 0
US-09-043-944-1 (1-461) x AF171064 (1-1335)

Qy 1 MetProSerThrArgArgGlnGlnGluGlyGlyAlaAspAlaGluThrHisThrVal 20
Db 1 ATGCCTTCCACAAGGAGACAACAGAGGGCGGAGGTGCAGATGCGGAAACACATACCGTT 60
Qy 21 TyxGlyThrAsnLeuIleThrAsnArgAsnSerGlnGluAspGluAsnValValGluGlu 40
Db 61 TACGGGTACAAATCTGATAACAAATCGAATACCAAGAACGAGAAAGAAATGTTGTGGAAGAA 120
Qy 41 AlaGluLeuLysTyxGlyAlaSerHisValIleHisLeuPheValProValSerLeuLys 60
Db 121 GCGGAGCTGAATAACGAGCATCTCAGTTATTCATCTATTTGTGCCGGTGTCACTATGC 180
Qy 61 MetAlaLeuValPheThrMetAsnThrIleThrPheTyxSerGlnAsnGlyArg 80
Db 181 ATGGCTCTGGTTGTTTTACGATGAACAGATTACGTTTTATAGTCAAAACAATGAAGG 240
Qy 81 HisLeuLeuSerHisProPheValArgGluThrAspSerIleValGluLysGlyLeuMet 100
Db 241 CATTTACTATATACCTCTTTTTCGCGGAAACAGACAGTATCGTTTGAGAGGATTTGATG 300
Qy 101 SerLeuGlyAsnAlaLeuValMetLeuCysValValValLeuMetThrValLeuLeuIle 120
Db 301 TCACCTTGGAAATGCTCTCGTCATGTTGCGTGGTCTGTCATGACAGTTCCTGCTGATT 360
Qy 121 ValPheTyxLysTyxPheTyxLysLeuIleHisGlyTyrLeuIleValSerSerPhe 140
Db 361 GTTTTCTATAAATAACAAGTTTTATAGCTTATTCATGGATGCTTATTGTCAGCAGTTTT 420
Qy 141 LeuLeuLeuPheLeuPheThrThrIleTyxValGlnGluValLeuLysSerPheAspVal 160
Db 421 CTTCCTCTCTTTCTTCTATTCACATAATCTATGTGCAAGAGTTCGAAAAAGTTTCGATGTG 480
Qy 161 SerProSerAlaLeuLeuValLeuPheGlyLeuGlyAsnTyxGlyValLeuGlyMetMet 180
Db 481 TCTCCACAGGCACTATTTGTTTGTGACTGGGTAACTATGGAGTTCCTCGGAATGATG 540
Qy 181 CysIleHisTyrLysGlyProLeuArgLeuGlnGlnPheTyxLeuIleThrMetSerAla 200
Db 541 TGTATACATTCGAAAGTCCATTGCGTCTGTGCAACAGTTCTACCTTATTACATGCTGCA 600
Qy 201 LeuMetAlaLeuValPheIleLysTyxLeuProGluTyrThrValTyrPheValLeuPhe 220
Db 601 CTAATGGCTCTGGTCTTTTCAAGTACCTACAGAAATGAGTGTGTGGTGTGTGCTGTTT 660
Qy 221 ValIleSerValTyrPheValAlaValLeuThrProLysGlyProLeuArgTyrLeu 240
Db 661 GTTATCTCGTTTGGATCTGGTTCGCGTGTCTACACCAAAAGGACCATTTGAGATATTG 720
Qy 241 ValGluThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyxSerSerGly 260
Db 721 GTGAAACATGCACAGGAGAGAAACGAGCAATTTTCCCGGCGCTGATTTATTTCGTCGA 780
Qy 261 ValIleTyxProTyxValLeuValThrAlaValGluAsnThrThrAspProArgGluPro 280
Db 781 GTCATCTATCCCTACGTTCTTTGTCTGCGGTGAAAAACACAGACAGACCCCGTGAACCG 840
Qy 281 ThrSerSerAspSerAsnThrSerThrAlaPheProGlyGluAlaSerCysSerSerGlu 300
Db 841 ACGTCGTCAGACTCAATACTTCTACAGCGTTTCTGAGAGGCGAGTGTTCATCTCTGAA 900
Qy 301 ThrProLysArgProLysValLysArgIleProGlnLysValGlnIleGluSerAsnThr 320
Db 901 ACGCAAAACGGCCAAAGTGAAACGAATTCCTCAAAAGTGCATAATCGAATCGAATACT 960
Qy 321 ThrAlaSerThrThrGlnAsnSerGlyValArgValGluArgGluLeuAlaGluArg 340

```

Db 961 ACAGTTCACAGACACAAACTCTGGAGTGAAGGTGGAAACGGAGCTAGCTGCTGAGAGA 1020

Qy 341 ProThrValGlnAspAlaasnPheHisArgHisGluGluGluArgGlyVallylsLeu 360
|||||

Db 1021 CCAACTGTACAGACGCCAATTTTCACAGGCACGAAGAGAGAGAGAGGTGTGAAACTT 1080

Qy 361 GlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSerSerTyrPheAsp 380
|||||

Db 1081 GGTCTGGGCGACTCATTTTCTACTCTGTCTCCCTGGCAAGGCTTCATCGTACTTTGAC 1140

Qy 381 TtpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGlyLeuCysPheThrLeuVal 400
|||||

Db 1141 TGGACACAGACTATCGCCTGTTATGGCCATCTTATCGGTCTCTGCTTCACTCTTGTC 1200

Qy 401 LeuLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPheProPheSerProAspSe 420
|||||

Db 1201 CTGCTCGCGCTCTCAACAGGACACTCCCGGCTCTGCCAATTTCCATTTCCTCGGACTC 1260

Qy 420 rPhePheThrPheValProAlaGlySerSerProHisLeuLeuHisLysSerLeuLysSe 440
|||||

Db 1261 ATTTTCTACTTTGTACCCGCTGGATCATCAACCCATTTGTTACACAACTCTCAAAAG 1320

Qy 440 rValTyrTyrIle 444
|||||

Db 1321 TGTATTATATAT 1333

RESULT 3

LOCUS U41540 27102 bp DNA linear INV 31-MAY-2003

DEFINITION Caenorhabditis elegans cosmid F35H12, complete sequence.

ACCESSION U41540

VERSION U41540.1 GI:1109862

KEYWORDS HTG.

ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 27102)
Wilson.R.

REFERENCE 1
AUTHORS Nelson,J. and Gattung,S.
TITLE The sequence of C. elegans cosmid F35H12
JOURNAL Unpublished (2001)

REFERENCE 2
AUTHORS Nelson,J. and Gattung,S.
TITLE The sequence of C. elegans cosmid F35H12
JOURNAL Unpublished (2001)

REFERENCE 3
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE 4
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE 5
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE 6
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE 7
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

TITLE JOURNAL

REFERENCE AUTHORS JOURNAL

REFERENCE AUTHORS JOURNAL

COMMENT

FEATURES source

gene

CDS

Direct Submission
Submitted (19-NOV-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

8 (bases 1 to 27102)
Waterston,R.

Direct Submission
Submitted (19-APR-2003) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

9 (bases 1 to 27102)
Wilson,R.

Direct Submission
Submitted (31-MAY-2003) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: submissions@watson.wustl.edu and jesus@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

For a graphical representation of this cosmid sequence and its analysis see:
<http://www.wormbase.org/db/seq/sequence?name=F35H12;class=Sequence>

NEIGHBORING COSMID INFORMATION

The 5' cosmid is ZC13, 200 bp overlap; the 3' cosmid is F53H8, 200 bp overlap. Actual start of this cosmid is at base position 1 of F35H12; actual end is at 12696 of F53H8.

NOTES:

Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program Genefinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yuji Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C. elegans ORFome cloning project (<http://worldb.dfci.harvard.edu/>), similarity to other proteins from BlastX analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

Location/Qualifiers
1..27102
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="X"
/clone="F35H12"
2463..3251
/gene="F35H12.6"
join(2463..2552,3063..3251)
/gene="F35H12.6"

```

/standard name="F35H12.6"
/notes="contains similarity to Homo sapiens Glycophorin E
precursor Glycophorin E precursor;
ENSEMBL:ENSP00000305872; coded for by the following C.
elegans cDNAs: Yk1092f08.3"
/codon_start=1
/product="Hypothetical protein F35H12.6"
/protein_id="AAM54184.1"
/db_xref="GI:21450554"
/db_xref="WormBase:F35H12.6"
/translation="MESYSGQVSESDSYSHESYSLNTVQGLAGVPMWAIWSFVICAV
LIIAVLLDVCVIRRNALGNTCTKARQKRGTINAAHQKRSSTNMLNV"
5010. .7295
/gene="sel-12"
Join(5010..5063,5134..5324,5629..5838,5884..6203,
6265..6348,6398..6604,7027..7295)
/gene="sel-12"
/standard name="F35H12.3"
/notes="contains similarity to Pfam domain PF01080
(Presenilin); coded for by the following C. elegans cDNAs:
Yk216el.3, Yk221d3.3, Yk221d3.5, Yk231a7.3, Yk231a7.5,
Yk400e8.3, Yk400e8.5, Yk452b9.3, Yk452b9.5, Yk499e3.3,
Yk573d4.3, Yk600e12.3, Yk674e3.3, AF171064, U35660,
OSTF019E11.1, OSTRO19E11.1"
/codon_start=1
/product="Suppressor/enhancer of lin-12 protein 12"
/protein_id="AAK39230.1"
/db_xref="GI:13775416"
/db_xref="WormBase:F35H12.3"
/translation="MPSTRROQEGGADAETHVYGTNLITNRNSQEDENVVEAEELK
YGASHVHLFVPSVLCMALVFMTNITFYSONGRHLLYTPFRETDSIVKGLMSL
GNALVMLVMTLVTLLVYKYKFKLHGWLIVSSFLLEFLTIIYQVQLKSFV
SPSALLVFLGNYGVGLNMCWHKGPRLRQOPLYLITMSALMALPIKPLPEWTVFV
LFVISWDLVAVLTPKGPURDYETAQRNEPFIALLYSGVVIYPLVLTAVENTD
PREPTSSNSNTAFPGASCSSETPKRPVKRIIPQKQIESNTASTTSGSVRVER
BLAAPTQVQDANFHRHBEERGVKGLGDFIYFVLLGKASSYFDMWNTTIACYVAIL
IGLCFTLLVLAVKRALPALPISIFSLGFIYFCTRWIITPFTVQVSQKLLY"
7533. .9207
/gene="F35H12.5"
Join(7533..7580,7730..7867,7922..8227,8270..8349,
8659..8744,8792..9207)
/gene="F35H12.5"
/standard name="F35H12.5a"
/notes="contains similarity to Interpro domains IPR000734
(Lipase), IPR000379 (Esterase/lipase/thioesterase, active
site); coded for by the following C. elegans cDNAs:
Yk703h10.5"
/codon_start=1
/product="Hypothetical protein F35H12.5a"
/protein_id="AAK39231.2"
/db_xref="GI:16950403"
/db_xref="WormBase:F35H12.5a"
/translation="MFSRRHILSNVRPIIKNLDLTRGFSDRKTYFDLNDLITKNVVT
YKTVNLEBETTVESKYVDVTPSSNTKGTIVTLSGPGTHNDFKYMKSFFQKKIRLI
CTNYPGSEFVGGHNSYTNQDRNSYKMSLMELELKNVRLIIMGHSRGGENALQLT
SMLNDENWPLGVAVINSPGAPHKGISKMGTFNFIISLKRHNKINSILPILH
YFYNLIGLRYSHGKVAALILPMQTFAPDEQKLSIDDLRAKPGIRAFYGYGSKDFLI
DEHQSSEVAMTFSESDHVTISNKKQDAEKAIKRSKSTTGQFVTANFKEGHLQKT
YFEFIVEVDSIFDADKEVDTKI"
Join(8003..8220,8269..8349,8659..8744,8792..9207)
/gene="F35H12.5"
/standard name="F35H12.5b"
/notes="contains similarity to Interpro domains IPR000734
(Lipase), IPR000379 (Esterase/lipase/thioesterase, active
site); coded for by the following C. elegans cDNAs:
OSTF212H6.1, OSTF212H6.1"
/codon_start=1
/product="Hypothetical protein F35H12.5b"
/protein_id="AAP46284.1"
/db_xref="GI:31249887"
/db_xref="WormBase:F35H12.5b"
/translation="MKSPFEQKIRLICTNYPGSEFVGGHNSYTNQDRNSYKMSLM
ETLELKNVRLIIMGHSRGGENALQLTSMNDENWPLGVAVINSPGAPHKGISKRMG

```

```

TINFIISLKRHNKNTINSILPHILHYFNLLIGLRVSHGKVAALILPMQTFAPDEQK
LSIDDLRAKPGIRAFYGYGSKDFLIIDEHQSSEVAMTFSESDHVTISNKKQDAEKAIK
RKSFTTGQFVTANFKEGHLQKTYPEFIVEVDSIFDADKEVDTKI"
complement(9443..113031)
/gene="F35H12.4"
complement(Join(9443..9627,10392..10517,10573..10772,
10828..11054,11147..11335,11383..11511,11615..11845,
11895..12010,12526..12728,12776..12893,12941..13031))
/gene="F35H12.4"

Alignment Scores:
Pred. No.: 3 33e-170 Length: 27102
Score: 2052.50 Matches: 459
Percent Similarity: 58.82% Conservative: 1
Best Local Similarity: 58.70% Mismatches: 1
Query Match: 86.31% Indels: 322
DB: 3 Gaps: 7

US-09-043-944-1 (1-461) x U41540 (1-27102)
QY 1 MetProSerThrArgArgGlnGlnGluGlyGlyGlyAlaAspAlaGluThrHis----- 18
Db 5010 ATGCCTTCCACAGGAGACACACAGGAGGGGGAGGTGCAGATCGCGAAGACACATGAAGT 5069
QY 18 ----- 18
Db 5070 TATTTAGACATTTATTTTCTCAAGAACTAAATGTTAAAAATGCTACAATGCATTTGTT 5129
QY 19 ----ThrValTyrGlyThrAsnLeuIleThrAsnArgAsnSerGlnGluAspGluAsnVa 37
Db 5130 TCAGACCGTTTACGGTACAAATCTGATACAATCGGAATAGCAAGAACGAAAAATGT 5189
QY 37 lValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIleHisLeuPheValProVa 57
Db 5190 TGTGGAGAGCGGAGCTGAATACGAGCATCTCAGTTATTCATCTATTGTGCGCGT 5249
QY 57 lSerLeuCysMetAlaLeuValValPheThrMetAsnThrIleThrPheTyrSerGlnAs 77
Db 5250 GTCACTATGCATGGCTCTGGTTGTTTACGATGAACACAGATTACGTTTATAGTCAAA 5309
QY 77 PAsnGlyArgHisLeu-----LeuSer----- 82
Db 5310 CAATGGAAGCAATTGTGAAGTTTCTAAGAANTGATTGAATAAATATTGATTTGTT 5369
QY 83 ----- 84
Db 5370 TTATCAATTTGCATCTGTGCATCGCACTCTTTGTGTCAGTCAAAATAATTTTGGTCAGT 5429
QY 84 ----- 84
Db 5430 GCAAAATAATTTTGTGTCAGTCAATGGTATTATGTGTGTCAGTCAATTTGCAAGTCTGAGT 5489
QY 84 ----- 84
Db 5490 TTAATACTTTTCGTGGTGTAAATTTTACTCAATTTTCTCAATATTCTTTTGGAAAAA 5549
QY 84 ----- 84
Db 5550 GTTGAAGATTACTCTGGAATTTTCCGAATAAATGTTAAATGGAATAATCAATCAACACA 5609
QY 85 -----His-ProPheValArgGluThrAspSerIleValG 96
Db 5610 AACTTTGAAATTTATTTTCAGACTATACACTCTCTTTTGTCCGGGAACACAGATGATCGTTG 5669
QY 96 lLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysValValValLeuMetT 116
Db 5670 AGAAGGGATTGATGTCACTTCGAAATGCTCTCGTCATGTTTGTGCGTGTCTCTCTGATGA 5729
QY 116 hrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuIleHisGlyTrpLeuI 136
Db 5730 CAGTTCTGCTGATTTGTTCTTATTAATACAGTTTATTAAGCTATTATTCATGATGGCTTA 5789
QY 136 leValSerSerPheLeuLeuPheLeuPheThrThrIleTyrValGln----- 152

```

[illegible]

D	b		6869	AGATTCAAGAAATCTCAATAAACAACCTTGCAGCTGCCACTTCGGCGTAGTCACGAAAAA	6929
Q	y		356	-----	356
D	b		6929	AAGGATTTCTTCAATAACCCTCTGTTCATATTTTTTTTAACAATAATTTTTCATCTCTC	6988
Q	y		357	-----	357
D	b		6989	ATATTTTGATATGTTTTTGCACAAAAAATGATTGCAGGAGGTGTGAACACTTGTGCTGGG	7048
Q	y		363	YAspPheIlePheTySerValLeuLeuGlyLysAlaSerSerTyrPheAspTrpAsnTh	383
D	b		7049	CGACTTCATTTTCTACTCTGTTCTCCTCGCAAGGCTTCATCGTACTTTGACTTGAACAC	7108
Q	y		383	rThrIleAlaCysTyrValAlaIleLeuIleGlyLeuCysPheThrLeuValLeuLeuAl	403
D	b		7109	GACTATCGCTTGATGTGGCAATCTTATCGGTCTCTGCTTCACTCTTGTCTCTCGCTCGC	7168
Q	y		403	aValPheLyArgAlaLeuProAlaLeu-GlnPheProPheSerProAspSerPhePheT	423
D	b		7169	CGTCTTCAAACGAGCACTCCCGGCTCTGCCAATTTCATTTTCTCCGGACTCATTTTTTA	7238
Q	y		423	hrPheValProAlaGlySerSerProHisLeuLeuHisLysSerLeuLysSerValTyT	443
D	b		7229	CTTTGTACCCGCTGGATCATCACCCCATTGTTACAGAATCTCTCACAAGTGTATTATT	7288
Q	y		443	YrIleAsnSerLeuPheLeuProPheLeuCysIlelleAsnPheSerIlelleSer	461
D	b		7289	ATATTAATTCITGTTTTTGGCAATTTCTTTGCATCATCAACTTTTCGATATATCT	7344
R	E	S	U		
RESULT	4				
AY043492					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
CDS					

LVKASATASGDMNTTACFAVALLIGLCILTLALLLALFKKALPALPISITIFGLVFPAT
 DNLVPPMDQLAFQPII"
 BASE COUNT 502 a 389 c 406 g 529 t 10 others
 ORIGIN

Alignment Scores:

Pred. No.: 1 42e-79 Length: 1836
 Score: 1009.50 Matches: 227
 Percent Similarity: 65.10% Conservative: 77
 Best Local Similarity: 48.61% Mismatches: 127
 Query Match: 42.45% Indels: 36
 DB: 5 Gaps: 10

US-09-043-944-1 (1-461) x AY043492 (1-1836)

QY 3 SerThrArgArgGlnGlnGluGlyGlyAlaAspAlaGluThrHisThrValTyrGly 22
 DB 203 AGTGAGCGCGGAAGAATGACCAATCTGGTCTGAGACCAATGGGCGACCAACAAC 262
 QY 23 ThrAsnLeuLeuThrAsnArgAsnSerGlnGluAspGluAsnValValGluGluAlaGlu 42
 DB 263 ATTCAGCAAGTGGTGATCAAGATGAAGAAGAT-----GAGGAGCTAACA 310
 QY 43 LeuLysTyrGlyAlaSerHisValIleHisLeuPheValProValSerLeuCysMetAla 62
 DB 311 CTGAATATGGGCGCAAAACATGATGATTGTTGTTGTTACCTGTACATTTGCAATGGTG 370
 QY 63 LeuValValPheThrMetAsnThrIleThrPheTyrSerGlnAsnAsnGlyArgHisLeu 82
 DB 371 GTCGTGTGTGCAACCATCAAGTCTGTGACGCTTTATACAGCAAGGATGGACAG--CTT 427
 QY 83 LeuSerHisProPheValArgGluThrAspSerIleValGluLysGlyLeuMetSerLeu 102
 DB 428 AVCTACACCTCCCTTCACAGAGACAGACATCCATTGGACAGAGGCGCCCTGAATCAATT 487
 QY 103 GlyAsnAlaLeuValMetLeuCysValValValLeuMetThrValLeuLeuValPhe 122
 DB 488 CTCACAGCTGCAATCATGATTAGTGTCAATGTCATGACCATCATCTCGTGTGTGCTT 547
 QY 123 TyrLysTyrLysPheTyrLysLeuLeuHisGlyTyrLeuIleValSerSerPheLeuLeu 142
 DB 548 TACAAATACAGGTGTACAAAGGTATCCACGGTGGCTTATCATTTCTCTCTTTGCTG 607
 QY 143 LeuPheLeuPheThrThrIleTyrValGlnGluValLeuLysSerPheAspValSerPro 162
 DB 608 CTTTTCTTTTTCATCTATCTATGAGTGAGGTTTTTAAAGCGTATAACGCTGCCATG 667
 QY 163 SerAlaLeuLeuValPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMetCysIle 182
 DB 668 GACTACATTACAGTGGCATATCTGGAACCTTGGTGTGGGAATGATTGTATC 727
 QY 183 HisTyrLysGlyProLeuArgLeuGlnGlnPheTyrLeuIleThrMetSerAlaLeuMet 202
 DB 728 CATTGGAAGGGCCCACTCGGCTTCAGCAAGCATATCTCATATGATAAGCGCTCTCATG 787
 QY 203 AlaLeuValPheIleLysTyrLeuProGluTyrThrValTyrPheValLeuPheValIle 222
 DB 788 GCATTGGTGTTCATTAAAGTATCTCCGAATGGACTGATGCTTATCTGCTGCTCAT 847
 QY 223 SerValTyrAspLeuValAlaValLeuThrProLysGlyProLeuArgTyrLeuValGlu 242
 DB 848 TCAGTGACGATTGGTGTCTTATGCTCTTAAGGTCCTCTTCGTATGTATTGTGA 907
 QY 243 ThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSerSerGlyValIle 262
 DB 908 ACAGCTCANGAGAGAAATGAACCGCTCTCCAGCACATTATTTACTCTCAACATGGTG 967
 QY 263 Tyr-----ProTyrValLeuValThrAlaValGluAsnThr 274
 DB 968 TGGCTAGTGAAATCGGCTGAGGAAGATCCAGAGCCCAACGGAAAGCTTCTAAACATCC 1027
 QY 275 ThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPheProGlyGlu 294
 DB 1174

DB 1028 ACTTACGATAAACAAGCTCGGCAACACAGTCTCAGAATGAAGATGCTGAAGCAGATGAT 1087
 QY 295 AlaSerCysSerSerGluThrProLysArgProLysValIleArgIleProGlnLysVal 314
 DB 1088 GNGGGCTTTAGTCAAGAGTGGCAGCAACAAGAGACAATAAGATAGGACCC----- 1138
 QY 315 GlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArgValGluArg 334
 DB 1139 ---ATTGAATCA-----ACACCTGAATCAGGAGCTGCTGTTTCAG--- 1174
 QY 335 GluLeuAlaAlaGluArgProThrValGlnAspAlaAsnPheHisGluGluGlu 354
 DB 1175 GCTCTGCTAGTAACCTCAACACAGTAGAGAC-----CTCGAA 1213
 QY 355 GluArgGlyValLysLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLys 374
 DB 1214 GAAAGAGGAGTAAAGCTAGTGTTAGGAGACTTCATTTCTACAGCGCTCTGGTGGCAA 1273
 QY 375 AlaSerSerTyrPhe-----AspTyrAsnThrThrIleAlaCysTyrValAlaIleLeu 392
 DB 1274 GCCTCTGCAACGGCAAGCGGGATTGGAATACAACCTTTAGCATGTTTGTAGCCATATTA 1333
 QY 393 IleGlyLeuCysPheThrLeuValLeuAlaValPheLysArgAlaLeuProAlaLeu 412
 DB 1334 ATGGTCTGTGCTTACACTCTGCTCTGTGCTCTTTAAGAGGCGCTTACAGCTCTT 1393
 QY 413 -GlnPheProPheSerProAspSerPhePheValProAlaGlySerSerProHi 432
 DB 1394 CCAATCTCCATTACTTTGGGCTAGTTTACTTTGCTACAGATA---ACTGGTGCAG 1450
 QY 432 sleuLeuHisLysSerLeuLysSerValTyrTyrIleAsnSer---LeuPheLeuProPh 451
 DB 1451 CCTTTATGGACGACGCTAGCATTCATCATGTTTATATCATGACGCTTGGTGTCTGATAT 1510
 QY 451 eLeuCysIleIleAsnPhe 457
 DB 1511 TTTATGATATATTGGATTT 1529
 RESULT 5
 AF038936
 LOCUS Bos taurus presenilin 1 mRNA, complete cds.
 DEFINITION
 ACCESSION AF038936
 VERSION AF038936.1 GI:5070115
 KEYWORDS
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 1896)
 AUTHORS Sahara.N., Shirasawa.T. and Mori.H.
 JOURNAL Molecular cloning of bovine presenilin 1 gene
 UNPUBLISHED
 REFERENCE 2 (bases 1 to 1896)
 AUTHORS Sahara.N., Shirasawa.T. and Mori.H.
 TITLE Direct Submission
 JOURNAL Submitted (16-DEC-1997) Molecular Biology, Tokyo Institute of
 Psychiatry, 2-1-8 Kami-Kitazawa, Setagaya-ku, Tokyo 156, Japan
 FEATURES
 location/Qualifiers
 1..1896
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="brain"
 9..1445
 /note="causal gene for familial Alzheimer's disease"
 /codon_start=1
 /product="presenilin 1"
 /protein_id="AAD39023.1"
 /db_xref="GI:5070116"
 /translation="MTSLPAPLSYFQNAQMSDNHLSNTVRSQNSDRERHGHGNRRR
 RGNTEVSNGRAPSSQQVVEQEEDEELTKYGAKEHIMLFPVTLMMVVAVATIK

795	GGGAGAGTGTAAACCTATAACGTTCTGTGGACTACATTA	CTGTGTGCACTCCTCATC	854	
172	GlyAsnTyrGlyValLeuGlyMetMetCysIleHis	trpLysGlyProLeuArgLeuGln	191	
855	TGGAATTTGGTGTGGGAATGATTTCCATTCACTGGAAGG	TCCACTTCGACTCCAG	914	
192	GlnPheTyrLeuIleThrMetSerAlaLeuMetalala	ValPheIleLysTyrLeuPro	211	
915	CAGGCATATCTCATTAATAGTAGGCCCTCATGGCCCT	TGGTGTATCAAGTACTCCTCC	974	
212	GluTrpThrValTrpPheValLeuPheValIleSer	ValTrpAspLeuValAlaValLeu	231	
975	GAATGGACTCGTGCCTCATCTTGGCTGTGATTTCC	AGTATATAGTGGCTGTTTG	1034	
232	ThrProLysGlyProLeuArgTyrLeuValGluThr	AlaGlnGluArgAsnGluProIle	251	
1035	TGTCGGAAGGTCACCTTCGTATCGTGTGGATTGAA	ACAGCTCAGGAGAGAAATGAAACGCTT	1094	
252	PheProAlaLeuIleTyrSerSerGlyValIleTyr	ProTyrValLeuValThrAlaVal	271	
1095	TTTCCAGCTCTCATTTACTCTCTCAACATGGTGG	-----TTGTGGAATATGGCA	1145	
272	GluAsnThrThrAspProArgGluProThrSerSer	AspSerAsnThrSerThrAlaPhe	291	
1146	GAAGGAGACCCGGAAGCTCAAGAGGAGTAGTCCA	AAAAATTC-----	1187	
292	ProGlyGluAlaSerCysSerSerGluThrProLys	ArgProLysValLysArgIlePro	311	
1188	-----AGTATAATGCGAAGCACAGAAAG-----	-----	1214	
312	GlnLysValGlnIleGluSerAsnThrThrAlaSer	ThrThrGlnAsnSerGlyValArg	331	
1215	-----GAGTCACAAGACACTGTTGCAGAGATG	ATGATGCGGGTTCAGT	1259	
332	ValGlu-----	-----ArgGluLeuAlaAlaGlu	339	
1260	GAGGAATGGGAAGCCACAGAGGACAGTCACTAG	GGCTCATCGCTCTACACCTGAGTCA	1319	
340	ArgProThrValGlnAspAlaAsn-----	-----PheHisArgHisGluGluGluArg	356	
1320	CGAGCTGCTGTCCAGGAACTTCCAGCAGTATCCT	CGCTGGTGAAGACCCAGGAGAAAGG	1379	
357	GlyValLysLeuGlyLeuGlyAspPheIlePheTyr	SerValLeuLeuGlyLysAlaSer	376	
1380	GGAGTAAACTTGGATTGGGAGATTTCATTTTTC	PACAGTGTCTCGTTGGTAAAGCCTCA	1439	
377	SerTyrPhe-----	-----AspTrpAsnThrThrIleAlaCysTyr	ValAlaIleLeuIleGly	394
1440	GCAACACCCAGTGGAGACTGGAACACAAACATAG	CCCTGTTTCGTPAGCCATTAATAATGGT	1499	
395	LeuCysPheThrLeuValLeuAlaValPheLysArg	AlaLeuProAlaLeuGlnPhe	414	
1500	TTGTGCCTTACATTAATTAATTAATTAATTAAT	TAATTAATTAATTAATTAATTAATTAATTAAT	1559	
414	eProPheSerProAspSerPhePheThrPheVal	ProAlaGlySerSerProHisLeuLe	434	
1560	TCCA-----TCACCTTTGGCTGTGTTTCTACT	TTTGCCACAGATTATCTGTACAGCCTTTT	1616	
434	uHisLysSerLeuLysSerValTyrTyrIleAsn	SerLeuPhe-----LeuProPhe	451	
1617	ATGGACCAATTAGCATTCATCAATTTATATCTAG	CATATTTGCGGTGAGATCCCATG	1676	
451	eLeuCysIleIle	455		
1677	GATGTTTCTTCT	1689		

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Hale,R.S., Rowley,A.G. and Blackstock,W.G.
TITLE Method of identifying modulators of presenilin
JOURNAL Patent: WO 0167109-A 5 13-SEP-2001;
GLAXO GROUP LIMITED (GB)
FEATURES Location/Qualifiers
source 1..2765
 /organism= "Homo sapiens"
 /mol_type= "genomic DNA"
 /db_xref= "taxon.9606"
 249..1652
 /note= "unnamed protein product"
 /codon_start= 1
 /protein_id= "CAC88544.1"
 /db_xref= "GI:15859209"
 /translation= "MTLEPAPLSYFQNAQMSDENHLSNTVRSQNDNRERQEHNDRSL
 GHPEPLSNRGPQNSQVVQDEEDEELTKYGAKHIMLFVPVTLCMVAVATIKS
 VSRYTRKDGQLIYTPETETETVGQRALHSILNAAIMISIVVMITLLVLVKYCIK
 VIRAWLISLLGLFFSFYLGEVEFTKNVAVDIYTVALLIWNVFGVGMISHWKGP
 LRIOQAYLIMSALMAFLIKYLPENTAWLILAVISYDLVAVLCFKGFLRLMLETAG
 ERNETIFPALIYSTVMVFNMAEGPEQAQRKSNKYNAESTERESQDTVAENDDG
 GFSENEAQRDSHLGPHRSTPEERAADVQLSSILAGEDPEERGKVLGGDFIFYSVL
 VGKASATASGDWNVTTCACFAVILIGLICLTILLALLAIKKALPALPISITGLVFYFADF
 YLQVPMWDQLAFHQFYI"

BASE COUNT	715 y	624 c	652 g	773 t	1 others
ORIGIN					
Alignment Scores:					
Pred. No.:	6.62e-78	Length:	2765		
Score:	993.50	Matches:	216		
Percent Similarity:	65.84%	Conservative:	77		
Best Local Similarity:	48.54%	Mismatches:	104		
Query Match:	41.78%	Indels:	48		
DB:	6	Gaps:	9		
US-09-043-944-1 (1-461) x AX244146 (1-2765)					
Qy	32	GlnGlAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIle	51		
Db	438	GAGCAAGATGAGGAAGAAGATGAGGAGCTGCACATTGAAATATGCCGCCAAGCATGTGATC	497		
Qy	52	HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrlie	71		
Db	498	ATGCTCTTTTGCCCTGTGACTCTCTCGACTGGTGTGCTGTGCTACCAATAAGTCAGTC	557		
Qy	72	ThrPheTyrSerGlnAsnGlyArgHisLeuLeuSerHisProPheValargGluThr	91		
Db	558	AGCTTTTATACCCGGAAGATGGGCAG---CTAATCTATACCCCATTCACAGAAGATACC	614		
Qy	92	AspSerIleValGlnGlyGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal	111		
Db	615	GAGACTGTGGCCAGAGAGCCCTGCACCTCAATTCTGATGTCTGCTCATCATGATCAGTGTC	674		
Qy	112	valValLeuMetThrValleuLeuIleValPheTyrLysTyrLysPheTyrLysLeulle	131		
Db	675	ATTGTTGTGATGACTATCTCTCTCTGGTGTCTGTATAAATACAGGTGCTATAAGGTCATC	734		
Qy	132	HisGlyTrpLeuIleValSerSerPheLeuLeuPheLeuPheThrThrIleTyrVal	151		
Db	735	CATGCTTGCGCTTATTATATCATCTCTATTGTTGCTGTCTTTTTTTCATTACTTACTGTG	794		
Qy	152	GlnGluValLeuLysSerPheAspValSerProSerAlaLeuValLeuPheGlyLeu	171		
Db	795	GGCGAAGTGTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTCGATC	854		
Qy	172	GlyAsnTyrGlyValLeuGlyMetMetCysIleHisIleTrpLysGlyProLeuArgLeuGln	191		
Db	855	TGGAATTTTGGTGTGGTGGGAATGATTTCCTACTGCTGAAAGTTCCTACTGATCCAG	914		

Qy	152	GlnGluValLeuLysSerPheAspValSerProSerAlaLeuValLeuPheGlyLeu	171
Db	795	GGGGAAGTGTATAAACCTATAACGTTGCTGGACTACATTACTGTGCACCTCCTGATC	854
Qy	172	GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTrpLysGlyProLeuArgLeuGln	191
Db	855	TGGAAATTTTGGTGTGGGGAATGATTTCCATTCACTGGAAAGTCCACATTCGACTCCAG	914
Qy	192	GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro	211
Db	915	CAGGCATATCTCAATTAGTGCCTCATGCGCTGGTGTATTATCAAGTACCTCCCT	974
Qy	212	GluTrpThrValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeu	231
Db	975	GAATGGACTGCGTGGCTCATCTTCGCTGTGTATTCAGTATATGATTTAGTGGCTGTTTTG	1034
Qy	232	ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle	251
Db	1035	TGTCGGAAGGTCCACITTCGTATCTGTTGTAACAGCTCAGGAGAGAAATGAACCGCTT	1094
Qy	252	PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal	271
Db	1095	TTTCCAGCTCTCAATTACTCTCAACAATGGTGG-----TTGGTGAATATGGCA	1145
Qy	272	GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe	291
Db	1146	GAAGGAGACCGGAAGCTCAAGAGGAGATATCCAAAATTTCC-----1187	
Qy	292	ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro	311
Db	1188	-----AAGTAAATGCGAAGGACACAAAG-----1214	
Qy	312	GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg	331
Db	1215	-----GAGTCACAGACACTGTTGCGAGAGATGATGCGGGTTCAGT	1259
Qy	332	ValGlu-----ArgGluLeuAlaAlaGlu	339
Db	1260	GAGGAATGGGAAGGCCAGAGGGACAGTCATCTAGGGCTCATCGCTCTACACCTGAGTCA	1319
Qy	340	ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg	356
Db	1320	CGAGCTGCTGCCAGGAACCTTCCAGCGATCTCTCGCTGGTGAAGACCCAGAGAAAGG	1379
Qy	357	GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer	376
Db	1380	GGAGTAAACTTGGATTGGGAGATTCAATTTCTACAGTGTCTTGGTGTGAAGCCTCA	1439
Qy	377	SerTyrPhe-----AspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly	394
Db	1440	GCAACAGCCAGTGGAGACTGGAACACAAACCATAGCTGTTTCTGTTAGCCATTAATTGGT	1499
Qy	395	LeuCysPheThrLeuValLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPhe	414
Db	1500	TTTGCCTTCAATTATTACTCTTGCATTTTCAAGAAAGCATTCGCAGCTCTTCCAATC	1559
Qy	414	eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLe	434
Db	1560	TCCA---TCACCTTTGGCTGTGTTTCTACTTTCACAGATATCTTGTACAGCCCTTTT	1616
Qy	434	uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh	451
Db	1617	ATGACCAATTAGCATTCATCAATTTTATATCTAGCATATTTGCGGTGTAGAATCCCATG	1676
Qy	451	eLeuCysIleIle 455	
Db	1677	GATGTTTCTTCTT 1689	
RESULT 13			
AR087309 2791 bp DNA linear PAT 07-SEP-2000			
LOCUS AR087309			
DEFINITION Sequence 133 from patent US 5986054.			

Search completed: December 11, 2003, 07:41:27
Job time : 4266 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2003, 06:24:06 ; Search time 2279 Seconds
(without alignments)
4916.351 Million cell updates/sec

Title: US-09-043-944-1
Perfect score: 2378
Sequence: 1 MPSTRRQEGGGAETHV.....YYINSLFLPFLCINFSIIS 461

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_P/US09043944/runat_11122003_062400_3685/app_query.fasta_1.647
-DB=EST -QFWT=Fastap -SUPFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09043944 @CGN 1_1349 @runat_11122003_062400_3685 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	986	41.5	589	14	CB385197
2	944	33.7	562	14	CB391622
C 3	942	33.6	558	14	CB404201
C 4	834	35.1	494	14	CB395382
5	788	33.1	477	12	BI174328
C 6	734.5	30.9	884	14	CD513489
7	734	30.9	964	13	EX426340
8	724.5	30.5	1101	12	BM548100
9	713.5	30.0	1998	11	AK014706
10	702.5	29.5	661	9	AI675803
11	699	29.4	697	13	BU614355
12	698.5	29.4	703	9	AI925372
C 13	688.5	29.0	974	13	EX426339
14	681.5	28.7	680	9	AI831581
15	671	28.2	881	14	CD358381
16	665.5	28.0	797	13	BU847845
C 17	645	27.1	386	12	BI175258
18	644	27.1	943	14	CD512951
19	631	26.5	757	10	AW850572
20	631	26.5	1112	12	BM476222
21	628	26.4	735	10	AW850636
22	627.5	26.4	705	9	AW604230
23	627	26.4	733	10	AW850631
24	625	26.3	671	10	AW850640
25	625	26.3	714	9	AW604310
26	623.5	26.2	874	14	CA477482
27	623	26.2	1201	9	AL553513
28	621	26.1	375	14	CA7805
29	620	26.1	360	9	AV190680
30	620	26.1	640	9	AW604295
31	614	25.8	980	14	BY714205
32	612	25.7	1201	9	AL553084
33	611.5	25.7	601	9	AW604307
34	610.5	25.7	613	9	AW367465
35	609.5	25.6	581	10	AW850712
36	609	25.6	374	14	C66636
37	609	25.6	704	10	EG633935
38	608.5	25.6	587	9	AA993681
39	607.5	25.5	591	10	AW850561
40	607.5	25.5	592	9	AW604313
41	607.5	25.5	599	9	AW604231
42	606.5	25.5	937	14	CA496597
43	602.5	25.3	944	12	BI453842
44	600.5	25.3	593	10	AW850573
C 45	600.5	25.3	737	10	AW850632

ALIGNMENTS

RESULT 1
LOCUS CB385197
DEFINITION CB385197_2 AD-wrmcDNA Caenorhabditis elegans cdna, mRNA sequence.
ACCESSION CB385197
VERSION CB385197.1
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 589)
AUTHORS Reboul,J., Vaglio,P., Rual,J.F., Lamesch,P., Martinez,M., Armstrong

C.M., Li.S., Jacotot, L., Bertin, N., Janky, R., Moore, T., Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S., Endress, G.A., Jenna, S., Chevet, E., Papasotiropoulos, V., Tollas, P.P., Ptacek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H., Doucette-Stamm, L., Hill, D.E. and Vidal, M.

C. elegans ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression

Nat. Genet., (2003) In press

Contact: Vidal M

Marc Vidal Laboratory

Dana Farber Cancer Institute

1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA

Tel: 617 632 5180

Fax: 617 632 5739

Email: Marc_Vidal@dfci.harvard.edu

Sequence tag of Gateway entry clones. The primers used were designed on the predicted protein encoding ORF. C. elegans ORFeome cloning project : Contact david_hill@dfci.harvard.edu or marc_vidal@dfci.harvard.edu

POLYA=No.

FEATURES

source

Location/Qualifiers

1..589

/organism="Caenorhabditis elegans"

/mol_type="mRNA"

/strain="N2"

/db_xref="taxon:6239"

/sex="Hermaphrodite and male"

/tissue_type="whole animal"

/dev_stage="mixed stage"

/clone_lib="AD-wrmcDNA"

/note="The AD-wrmcDNA library was generated with poly(A)+

RNA isolated from both hermaphrodite and male N2 worms of

all larval stages, embryos, adults and dauers and the

subsequent generation of cDNAs by poly(A) priming. The

cDNAs were cloned into pPC86"

BASE COUNT 157 a 106 c 137 g 189 t

ORIGIN

Alignment Scores:

Pred. No.: 3.63e-88 Length: 589

Score: 986.00 Matches: 193

Percent Similarity: 98.97% Conservative: 0

Best Local Similarity: 98.97% Mismatches: 2

Query Match: 41.46% Indels: 0

EB: 14 Gaps: 0

US-09-043-944-1 (1-461) x CB385197 (1-589)

Qy 8 GlnGluGlyGlyGlyAlaAspAlaGluThrHisThrValTyrGlyThrAsnLeuIleThr 27
Db 3 CAGGGGGGGGAGGTGCAGATGCGGAACACATACCGTTTACGGTACAAATCTGATACA 62
Qy 28 AsnArgAsnSerGlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAla 47
Db 63 AATCGGAATAGCCAGAGAGAGAAATGTTGTGGGAAGACGGAGCTGAATACGGAGCA 122
Qy 48 SerHisValIleHisLeuPheValProValSerLeuCysMetAlaLeuValValPheThr 67
Db 123 TCTCAGCTATTATCTATTTGTGCGGTGCATCATGATGCGGTCTGTTGTTTACG 182
Qy 68 MetAsnThrIleThrPheTyrSerGlnAsnAnGlyArgHisLeuLeuSerHisProPhe 87
Db 183 ATGACACAGATTACGTTTATAGTCAAAACATGGAAGGCATTTACTATACACTCTTTT 242
Qy 88 ValArgGluThrAspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuVal 107
Db 243 GTCCGGGAACACAGACAGTATGTTGAGAGGATGATGTCATCTGGAAATGCTCTGTC 302
Qy 108 MetLeuCysValValValLeuMetThrValLeuLeuLeuValPheTyrLysPhe 127
Db 303 ATGTTGTCGGTGGTCTGATGACAGTCTGCTGATTGTTTCTATAAATACAAGTTT 362

Qy 128 TyrLysLeuIleHisGlyTrpLeuIleValSerSerPheLeuLeuPheLeuPheThr 147
Db 363 TATAAGCTTATTCATGGATGGCTTATGTGCAGAGTTTCTTCTCTTTTCTTATTCAC 422
Qy 148 ThrIleTyrValGlnGluValLeuLysSerPheAspValSerProSerAlaLeuVal 167
Db 423 ACAATCTATGTGCAAGAAGTTCTGAAAGTTTCGATGTGTCTCCAGCGCACATATTGGTT 482
Qy 168 LeuPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMetCysIleHisTrpLysGlyPro 187
Db 493 TTGTTTGACATGGGTAACTATGAGAGTTCTCGAATAGATGTATATCATTTGGAAGTCCA 542
Qy 188 LeuArgLeuGlnGlnPheTyrLeuIleThrMetSerAlaLeuMet 202
Db 543 TTGCGTCTGCAACAGTTCTACCTTATTACAATGTCTGCACATAATG 587

RESULT 2

CB391622

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Caenorhabditis elegans

Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

; Rhabditidae; Pelodierinae; Caenorhabditis.

REFERENCE

AUTHORS

Reboul, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong, C.M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T., Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S., Endress, G.A., Jenna, S., Chevet, E., Papasotiropoulos, V., Tollas, P.P., Ptacek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H., Doucette-Stamm, L., Hill, D.E. and Vidal, M.

C. elegans ORFeome version 1.1: experimental verification of the

genome annotation and resource for proteome-scale protein

expression

Nat. Genet., (2003) In press

Contact: Vidal M

Marc Vidal Laboratory

Dana Farber Cancer Institute

1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA

Tel: 617 632 5180

Fax: 617 632 5739

Email: Marc_Vidal@dfci.harvard.edu

Sequence tag of Gateway entry clones. The primers used were

designed on the predicted protein encoding ORF. C. elegans ORFeome

cloning project : Contact david_hill@dfci.harvard.edu or

marc_vidal@dfci.harvard.edu

POLYA=No.

Location/Qualifiers

1..562

/organism="Caenorhabditis elegans"

/mol_type="mRNA"

/strain="N2"

/db_xref="taxon:6239"

/sex="Hermaphrodite and male"

/tissue_type="whole animal"

/dev_stage="mixed stage"

/clone_lib="AD-wrmcDNA"

/note="The AD-wrmcDNA library was generated with poly(A)+

RNA isolated from both hermaphrodite and male N2 worms of

all larval stages, embryos, adults and dauers and the

subsequent generation of cDNAs by poly(A) priming. The

cDNAs were cloned into pPC86"

BASE COUNT 150 a 99 c 133 g 180 t

ORIGIN

Alignment Scores:

Pred. No.: 5.37e-84 Length: 562

Score: 944.00 Matches: 184

Percent Similarity: 98.92% Conservative: 0

Best Local Similarity: 98.92% Mismatches: 2

Query Match: 39.70% Indels: 0
 DB: 14 Gaps: 0
 US-09-043-944-1 (1-461) x CB391622 (1-562)

QY 8 GlnGluGlyGlyGlyAlaHspAlaGluThrHisThrValTyGlyThrAsnLeuLeuThr 27
 DB 3 CAGGAGGCGGAGGTGCAGATCGGAAACACATACCGTTTACGGTACAAATCTGTATAACA 62
 QY 28 AsnArgAsnSerGlnGluAspGluAsnValValGluGluAlaGluLeuLeuTyGlyAla 47
 DB 63 AATCGGAATAGCCAGAACACCAAAATGTTGGGAAGACGGAGCTGAAATACGGAGCA 122
 QY 48 SerHisValIleHisLeuPheValProValSerLeuLeuCysMetAlaLeuValValPheThr 67
 DB 123 TCTCACGTTATTCACTATTGTCGGGTGTCACTATGTCATGGCTCTGGTGTGTTTACG 182
 QY 68 MetAsnThrIleThrPheTySerGlnAsnAnglyArgHisLeuLeuSerHisProPhe 87
 DB 183 ATGAACACGATTACGTTTATATAGTCAAAACAAATGGAAGCAATTTACTATACACTCCITTT 242
 QY 88 ValArgGluThrAspSerIleValGluLeuLysGlyLeuMetSerLeuGlyAsnAlaLeuVal 107
 DB 243 GTCCGGGAACACACAGATATCGTTGGAAGGATGTGATGTCATTGGAAATGCTCTCGTC 302
 QY 108 MetLeuCysValValValLeuMetThrValLeuLeuLeuIleValPheTyLeuTyLysPhe 127
 DB 303 ATGTTGTGGGTGCTGCTGATGACAGTCTGCTGATGTTTCTATAAATACCAAGTTT 362
 QY 128 TyrLysLeuIleHisGlyTrpLeuIleValSerSerPheLeuLeuLeuPheLeuPheThr 147
 DB 363 TATAAGCTTATTCATGATGGCTTATGTCACAGTTTCTTCTCTTCTTCTTCTTCTTCT 422
 QY 148 ThrIleTyValGlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuVal 167
 DB 423 ACAATCTATGTGCAAGAGTCTGAAAGTTTCGATGTGTCCTCCAGCGCACTATTGGTT 482
 QY 168 LeuPheGlyLeuGlyAsnTyGlyValLeuGlyMetMetCysIleHisTrpLysGlyPro 187
 DB 483 TTGTTGGACTGGTAACTATGAGTCTCGGAATGATGTATATACATTGGAAGGTCCA 542
 QY 188 LeuArgLeuGlnGlnPhe 193
 DB 543 TTGCGTCTGCACAGTTC 560

RESULT 3
 CB404201/c
 LOCUS
 DEFINITION OSTRO19E11_2 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
 ACCESSION CB404201
 VERSION CB404201.1 GI:30745928
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditiodea
 1 (bases 1 to 558)
 Rebol,J., Vaglio,P., Rual,J.F., Lamesch,P., Martinez,M., Armstrong
 C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T., Hudson
 J.R., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,
 Endress,G.A., Jenna,S., Chevret,E., Papasotiropoulos,V., Tollas,P.P.,
 Pracek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H.,
 Doucette-Stamm,L., Hill,D.E. and Vidal,M.
 C. elegans Orfome version 1.1: experimental verification of the
 genome annotation and resource for proteome-scale protein
 expression
 Nat. Genet., (2003) In press
 Contact: Vidal M
 Marc Vidal Laboratory
 Dana Farber Cancer Institute
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739

FEATURES
 Source
 Location/Qualifiers
 1..558
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /sex="Hermaphrodite and male"
 /tissue_type="whole animal"
 /dev_stage="mixed stage"
 /clone_lib="AD-wrmcDNA"
 /notes="The AD-wrmcDNA library was generated with poly(A)+
 RNA isolated from both hermaphrodite and male N2 worms of
 all larval stages, embryos, adults and dauers and the
 subsequent generation of cDNAs by poly(A) priming. The
 cDNAs were cloned into pPC86"
 BASE COUNT 153 a 118 c 149 g 138 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8,43e-84 Length: 558
 Score: 942.00 Matches: 185
 Percent Similarity: 99.46% Conservative: 0
 Best Local Similarity: 99.46% Mismatches: 0
 Query Match: 39.61% Indels: 1
 DB: 14 Gaps: 0

US-09-043-944-1 (1-461) x CB404201 (1-558)

QY 249 GluProIlePheProAlaLeuIleTySerSerGlyValIleTyProTyValLeuVal 268
 DB 557 GAGCAATTTTCCGGCGCTGATTTATTCCTCTGGAGTCATCATCCCTACGTTCTTGT 498
 QY 269 ThrAlaValGluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSer 288
 DB 497 ACTGCACTTGAACACACAGACAGACCCCGTGAACCGACGTCGTCAGACTCAAACTCTCT 438
 QY 289 ThrAlaPheProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLys 308
 DB 437 ACAGCTTTTCTGGAGAGGCGAGTTGTTTCATCTGAAACGCCAAACGCCCAAAAGTGAAA 378
 QY 309 ArgIleProGlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSer 328
 DB 377 CGAATTCCTCAAAAAGTGCATAATCGAATCGAATACAGCTTCAACGACACAAACTCT 318
 QY 329 GlyValArgValGluArgGluLeuAlaAlaGluArgProThrValGlnAspAlaAsnPhe 348
 DB 317 GGAGTAAGGGTGAACGGGGAGCTAGCTGTCGAGAGACCAACTGTACAAGACGCCAATTTT 258
 QY 349 HisArgHisGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPheIlePheTy 368
 DB 257 CACAGGCCACGAGGAGGAGAGAGAGGTGTGAAACTTGGTCTGGCGGACTTCATTTCTAC 198
 QY 369 SerValLeuLeuGlyLysAlaSerSerTyPheAspTrpAsnThrThrIleAlaCysTy 388
 DB 197 TCTGTTCTCTCGCAAGGCTTCATCGTACTTTGACTTGGAAACACGACTATCGTCTGTAT 138
 QY 389 ValAlaIleLeuIleGlyLeuCysPheThrLeuValLeuLeuAlaValPheLysArgAla 408
 DB 137 GTGGCCATTTTATTCGGTCTCTGCTTCACCTTGTCTGCTCCCTCTTCAACAGCA 78
 QY 409 LeuProAlaLeu-GlnPheProPheSerProAspSerPhePheThrPheValProAlaG 428
 DB 77 CTCGCGGCTCTGCCAATTTCCATTTCTCGGACTCATTTTTTACTTTTGTACCCGCTG 18
 QY 428 YSerSerProHisLeu 433
 DB 17 ATCATCACCCCATTTG 2

```

RESULT 4
CB395382/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
COMMENT

FEATURES
    source
        1. 494
            /organism="Caenorhabditis elegans"
            /mol_type="mRNA"
            /strain="N2"
            /db_xref="taxon:6239"
            /sex="Hermaphrodite and male"
            /tissue_type="whole animal"
            /dev_stage="mixed stage"
            /clone_lib="AD-wrmcDNA"
            /note="The AD-wrmcDNA library was generated with poly(A)+
            RNA isolated from both hermaphrodite and male N2 worms of
            all larval stages, embryos, adults and dauers and the
            subsequent generation of cDNAs by poly(A) priming. The
            cDNAs were cloned into pPC86"
            147 a 100 c 123 g 124 t

BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 494
Score: 834.00 Length: 494
Percent Similarity: 99.30% Matches: 164
Best Local Similarity: 99.30% Conservatives: 0
Query Match: 35.07% Mismatches: 0
DB: 14 Gaps: 0

US-09-043-944-1 (1-461) x CB395382 (1-494)

QY 289 ThrAlaPheProGlyGluAlaSerCysSerSerGluThrProLysArgProLysVallys 308
DB 494 ACAGCTTTTCTGGAGGCGAGTGTTCATCTGAAACGCCAAACGGCCAAAGTGAA 435
QY 309 ArgIleProGlyLysValGlnIleGluSerAsnThrThrAlaSerThrGlnAsnSer 328
DB 434 CGAATCTCTCAAAAAGTCGAATCGAATCGAATCTACAGCTTCACAGCACAAAAC 375
QY 329 GlyValArgValGluArgGluLeuAlaAlaGluArgProThrValGlnAspAlaAsnPhe 348

```

```

Db 374 GGAGTAAGGCTGGAAACGGGAGCTAGCTGTGAGAGACCAACTGTACAAAGACGCCAATTTT 315
QY 349 HisArgHisGluGluGluArgGlyValLysLeuGlyLeuGlyAspPheIlePheTyr 368
Db 314 CACAGCACGAGAGAGAGAGAGAGAGAGTGTGAACTTGGTCTGGGCGACTTCATTTTCTAC 255
QY 369 SerValLeuLeuGlyLysAlaSerSerTyrPheAspTrpAsnThrThrIleAlaCysTyr 388
Db 254 TCTGTTCTCTCGGCAAGGCTTCATCGTACTTGTGACTGGAAACAGCATATCGCTTGTAT 195
QY 389 ValAlaIleLeuIleGlyLeuCysPheThrLeuValLeuAlaValPheLysArgAla 408
Db 194 GTGGCCATCTTATCGGCTCTCTCTTCACTCTTGTCTGCTCGCCGCTTCTCAACAGCA 135
QY 409 LeuProAlaLeu-GlnPheProPheSerProAspSerPheThrPheValProAlaG 428
Db 134 CTCCTGGCTCTGCCAATTTCCATTTCTCGGACTCATTTTTCATTTTGTATCCCGCTGG 75
QY 428 ySerSerProHisLeuLeuHisLysSerLysSerValTyrTyrIleAsnSerLeuPh 448
Db 74 ATCATCACCCCATTTGTACACAAGTCTCTCAAAAGTGTATTATATTATTAATTTCTCTGT 15
QY 448 eLeuProPheLeu 452
Db 14 TTGCGCATTTCTT 2

RESULT 5
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMED
COMMENT

FEATURES
    Location/Qualifiers
        1. 477
            /organism="Caenorhabditis elegans"
            /mol_type="mRNA"
            /strain="N2"
            /db_xref="taxon:6239"
            /sex="Hermaphrodite and male"
            /tissue_type="whole animal"
            /dev_stage="mixed stage"
            /clone_lib="AD-wrmcDNA"
            /note="The AD-wrmcDNA library was generated with poly(A)+
            RNA isolated from both hermaphrodite and male N2 worms of
            all larval stages, embryos, adults and dauers and the

```

subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pPC86"

BASE COUNT 131 a 87 c 109 g 150 t
ORIGIN

Alignment Scores:
Pred. No.: 1,68e-68 Length: 477
Score: 788.00 Matches: 156
Percent Similarity: 98.73% Conservative: 0
Best Local Similarity: 98.73% Mismatches: 2
Query Match: 33.14% Indels: 0
DB: 12 Gaps: 0

US-09-043-944-1 (1-461) x B1174328 (1-477)

QY 8 GlnGluGlyGlyGlyAlaAspAlaGluThrHisThrValTyrGlyThrAsnLeuLeuPheThr 27
DB 3 CAGGAGGGCGGAGGTCAGATCGGAAACACATACCGTTTACGGTACAAATCTGATACCA 62

QY 28 AsnArgAsnSerGlnGluAspGluAsnValValGluGluAlaGluLeuLeuTyrGlyVala 47
DB 63 AATCGGAATAGCCAAAGAGACGAAATGTTGTGGAAGAGCGGAGCTGAATACGGAGCA 122

QY 48 SerHisValIleHisLeuPheValProValSerLeuCysMetAlaLeuValValPheThr 67
DB 123 TCTACGTTATTATCATCTATTGTGCGGTGCACATCATGCTGCTGGTTGTTTACG 182

QY 68 MetAsnThrIleThrPheTyrSerGlnAsnAenGlyArgHisLeuLeuSerHisProPhe 87
DB 193 ATGACACAGATTACGTTTATAGTCAAAACATGAGGAGCATTTACTATACATCTTTT 242

QY 88 ValArgGluThrAspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuVal 107
DB 243 GTCCGGAAACAGACAGATATCTTTCAGAGGAGGATGATGTCACCTTGGAAATGCTCTGTC 302

QY 108 MetLeuCysValValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPhe 127
DB 303 ATGTGTGCGTGTCTTCTGAGCAGATTCGCTGATGTTTCTTAAATACAGTTT 362

QY 128 TyrLysLeuIleHisGlyTrpLeuIleValSerSerPheLeuLeuLeuPheLeuPheThr 147
DB 363 TATAAGCTTATTATCATGGATGCTTATTGTCAGCAGTTTCTTCTTCTTCTTCTTCTT 422

QY 148 ThrIleTyrValIleGluValLeuLysSerPheAsnValSerProSerAlaLeu 165
DB 423 ACAATCTATGTCAAGAGTTCTGAAAGTTTCGATGTTCTCCCGGCGCACTA 476

RESULT 6
CD513489/c
LOCUS
DEFINITION CD513489 884 bp mRNA linear EST 06-JUN-2003
AGENCY AGENCOURT 14365780 NIH_MGC_179 Homo sapiens cDNA clone
IMAGE:30394561 5', mRNA sequence.
CD513489
CD513489.1 GI:31445207
EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 884)
NIH-MGC http://imgc.ncbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNI)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM473 row: e column: 02
High quality sequence stop: 652.

FEATURES

Location/Qualifiers
1..884
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30394561"
/tissue_type="Pituitary"
/lab_host="NIH_MGC_179"
/notes="Organ: brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2: NotI; Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.1 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."
BASE COUNT 218 a 247 c 259 g 160 t
ORIGIN

Alignment Scores:
Pred. No.: 9.09e-63 Length: 884
Score: 734.50 Matches: 164
Percent Similarity: 64.92% Conservative: 34
Best Local Similarity: 53.77% Mismatches: 53
Query Match: 30.89% Indels: 54
DB: 14 Gaps: 7

US-09-043-944-1 (1-461) x CD513489 (1-884)

QY 113 ValLeuMetThrValLeu-LeuIleValPheTyrLysTyrLysPheTyrLysLeuIleHi 132
DB 785 GTGGTTATGACCATTCCTTGGTGGTCTCTCAAGTACCGTGTCAAGTTCATCCA 726

QY 132 sGlyTrpLeuIleValSerSerPheLeuLeuLeuPheLeuPheThrThrIleTyrValGI 152
DB 725 TGGCTGGTTGATGATCTTCACTGATGCTGCTTCCCTTCCCTATATCTACCTTGG 666

QY 152 nGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeuGI 172
DB 665 GGAAGTGCTCAAGACCTACAATGTGCCATGAGTACTCCCCACCTCTTGTGCTGTCTG 606

QY 172 YAnTyrGlyValLeuGlyMetMetCysIleHisTrpLysGlyProLeuArgLeuGI 192
DB 605 GAACCTCGGGCGCAGTGGCGATGTGTGCTCACTTCCACTGGAGGGCCCTCTGTGTCTGACGA 546

QY 192 nPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGI 212
DB 545 GGCCTACCTCATCATGATCAGTGGCTCATGGCCCTAGTGTTCATCAAGTACCTCCAGA 486

QY 212 uTrpThrValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeuTh 232
DB 485 GTGGTCCGCGTGGGTCACTCTCGCGCCCATCTCTGTGTATGATCTCTGTGCTGTGTG 426

QY 232 rProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIlePh 252
DB 425 TCCCAAGGGCCCTCGAATGCTGGTAGAACTGCCAGGAGAGAAATGAGCCATATT 366

QY 252 eProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThr----- 289
DB 365 CCCTGCCCTGATATCTCATCTGCCATGGTGGTGGCTGGCATGGCGAAGTGGAGCC 306

QY 270 -----AlaValGluAsnThrThrAspProArg---GluProThrSerSerAs 284
DB 305 CTCCTCTCAGGGTGGCCCTCCAGCTCCCTACGACCCCGAGATGGAGAGACTTCCTATGA 246

QY 284 pSerAsnThrSerThrAlaPheProGlyGluAlaSerCysSerSerGluThrProLysAr 304
DB 245 CAGTTT-----GGGAG-----CCTTCATA 225

QY 304 gProLysValLysArgIleProGlnLysValGlnIleGluSerAsnThrThrAlaSerTh 324
DB 304 gProLysValLysArgIleProGlnLysValGlnIleGluSerAsnThrThrAlaSerTh 324

```

Db      224 CCCGAGAGCTTTGAGCCTCC-----TT 201
QY      324 rThrGlnAsnSerGlyValArgValGluArgGluLeuAlaAlaGluArgProThrValG1 344
Db      200 GACTGGCTACCCAGGGAGGAGCTGGAG-----173
QY      344 nAspAlaAsnPheHisArgHisGluGluGluGluArgGlyValLysLeuGlyLysAs 364
Db      172 -----GAAGAGAGAGGAAAGGGCGCTGAAGCTTGGCTCGGGGA 135
QY      364 pPheIlePheTyrSerValLeuLeuGlyLysAlaSerSer-----TyrPheAspTrpAs 382
Db      134 CTTTCATCTTCTACAGTGTGCTGGTGGGCAAGCGGCTGCCAGCGGAGGAGCTGGAA 75
QY      382 nThrThrIleAlaCysTyrValAlaIleLeuIleLeuGlyLeuGlyCysPheThrLeuValLeu 402
Db      74 TACCAGCTGGCTGCTGCTGGCCATCTCTATTGGCTGTGTCTGACCCCTCCTGCTGCT 15
QY      402 uAlaValPheLys 406
Db      14 TGCTGTGTTCAG 2

RESULT 7
BX426340
LOCUS      BX426340 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0AG004YO20 5-PRIME, mRNA sequence.
ACCESSION  BX426340
VERSION    BX426340.1 GI:30772565
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS    Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished
COMMENT    Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 595.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0AG004BH100P1&cluster=595.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0AG004BH100P1.
FEATURES   Location/Qualifiers
            1..964
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DG004YO20"
                /tissue_type="B CELLS (RAMOS CELL LINE)"
                /cell_line="RAMOS CELL LINE"
                /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                with a NotI-oligo(dT) primer. Five prime end enriched,
                double-strand cDNA was digested with Not I and cloned into
                the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                Library was not normalized."
BASE COUNT      195 a      266 c      271 g      230 t
ORIGIN
Alignment Scores:
Pred. No.:      1,16e-62      Length:      964
Score:          734.00      Matches:      162
Percent Similarity: 64.72%      Conservative: 49
Best Local Similarity: 49.69%      Mismatches: 69
Query Match:    30.87%      Indels:      47

```

RESULT 8

```

DB:      13      Gaps:      5
US-09-043-944-1 (1-461) x BX426340 (1-964)
QY      75 SerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThrAspSerIle 94
Db      105 TCCGAGCACTTTGGGAGACTGAGCATCTACAGCCATTTCATGAGGACACACCCCTCGGTG 164
QY      95 ValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuLeuMetLeuValValLeu 114
Db      165 GGCCAGCGCCTCTCAACTCCGTCGTGAACACCTCATCATCATCATCATCATCATCATCAT 224
QY      115 MetThrValLeuLeuValPheTyrLysTyrLysPheTyrLysLeuLeuHisGlyTyr 134
Db      225 ATGACCATCTTCTTGGTGTGCTCTACAAGTACCGCTGTCAAGTTCATCAAGTTCATCAAGT 284
QY      135 LeuIleValSerPheLeuLeuLeuPheLeuPheThrThrIleTyrValGlnGluVal 154
Db      285 TTGATCATGTCTTCACTGATGTCTGTCTCTTCACTATATCTACCTATATCTACCTTGGGGA 344
QY      155 LeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeuGlyAsnTyr 174
Db      345 CTCAGACCTACATGTGCCCATGACTACCCACCCCTCTTCTGACTCTCTGGAACCTTC 404
QY      175 GlyValLeuGlyMetMetCysIleHisTyrLysGlyProLeuArgLeuGlnGlnPheTyr 194
Db      405 GGGGAGTGGGCGATGCTGTCATCCATCGAAGGGCCCTCTGCTGTGTCGAGCAGGCTAC 464
QY      195 LeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluThrPThr 214
Db      465 CTCATCATGATCAGTCGCTCATGCCCTAGTGTTCATCAAGTACCTCTCCAGAGTGGTCC 524
QY      215 ValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeuThrProLys 234
Db      525 CGCTGGGTCTCTCTGGGCCCATCTCTGTGTATGATCTCTGTGGTGTCTGTGTCCCAA 584
QY      235 GlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIlePheProAla 254
Db      585 GGGCCTCTGAGAATGCTGTAGAAACTGCCAGGAGAGAAATGAGCCCATATTTCCTGTC 644
QY      255 LeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaValGluAsnThr 274
Db      645 CTGATATACTCATCTGCCATGCTGTGGACGGTGTGGCATGGCGAATGN-----692
QY      275 ThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPheProGlyGlu 294
Db      693 -----GACCCCTCTCTCAGGGT-----710
QY      295 AlaSerCysSerSerGluThrProLysArgProLysVal-LysArgIleProGlnLysVa 314
Db      711 -----GCCCTCCAGCTCCCTTACGCCGAGATGGAAGAAGACTCTCTATGACAGT 761
QY      314 IglNileGluSerAsnThrThrAlaSerThrThr-----GlnAsnSerGlyVa 330
Db      762 TTTGGGAGCCTTTCATACCCGAGCTTTTGAGCCTCCCTTGACTGGCTACCCAGGGGA 821
QY      330 IargValGluArgGluLeuAlaAlaGluArgProThrValGlnAspAlaAsnPheHisAr 350
Db      822 GGAGCTGGAG-----831
QY      350 gHisGluGluGluArgGlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerVa 370
Db      832 ---GAAGAGAGGAAAGGGCGTGAAGCTTGGCTCGGGGACTTCTCTTCTACAGTGT 887
QY      370 lLeuLeuGlyLysAlaSer-----SerTyrPheAspTrpAsnThrThrIleAlaCysTy 388
Db      888 GCTGTGGGCAAGCGGCTGCACGGGCACGGGG-GACTGGGAATACCCAGCTCGCTGCTT 946
QY      388 rValAlaIleLeuIle 393
Db      947 CGTGGCCATCTTCATT 962

```


TITLE
JOURNAL
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1. .697
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="UI-M-EV0-cbf-1-22-0-UI"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP EV0"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 , is GTGGCTGGAA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."
 BASE COUNT 160 a 173 c 158 g 204 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2-29e-59 Length: 697
 Score: 699.00 Matches: 135
 Percent Similarity: 77.83% Conservative: 44
 Best Local Similarity: 58.70% Mismatches: 47
 Query Match: 29.39% Indels: 4
 DB: 13 Gaps: 2

US-09-043-944-1 (1-461) x BU614355 (1-697)

QY 43 LeuLysTyrGlyAlaSerHisValIleHisLeuPheValProValSerLeuCysMetAla 62
 Db 7 TTGAATATGGAGCAAGCATGTCATCATGCTCTTTGTCCCGTGCACCTCTGCATGGTC 66
 QY 63 LeuValValPheThrMetAsnThrIleThrPheTyrSerGlnAsnAnglyArgHisLeu 82
 Db 67 GTGCTGTGGCCACCAATCAATCAGTCAGCTTCTATACCGGAGGACGCTCAG---CTA 123
 QY 83 LeuSerHisProPheValArgGluThrAspSerIleValGluLysGlyLeuMetSerLeu 102
 Db 124 ATCTACACCCATTCACAGNAGACACTGAGACTGTAGGCCAAGAGCCCTGCATCGATC 193
 QY 103 GlyAsnAlaLeuValMetLeuCysValValValLeuMetThrValLeuLeuValPhe 122
 Db 184 CTGAATGCGGCCATCATGATGATGTCATGTCATTATGACCATCTCCTCGTGGTGGCTG 243
 QY 123 TTTTysTyrLysPheTyrLysLeuIleHisGlyTyrLeuIleValSerPheLeuLeu 142
 Db 244 TATAAATACAGGTGTCACAGGTCATCCACGCTGCTTATTATTCTCTCTGTGGTGG 303
 QY 143 LeuPheLeuPheThrThrIleTyrValGlnGluValLeuLysSerPheAspValSerPro 162

Db 304 CTGTTCTTTTTCGTTTCATTTAGGGAAGTAGTTTAAAGACCTACAAATGTCGCGTG 363
 QY 163 SerAlaLeuLeuValLeuPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMetCysIle 182
 Db 364 GACTAGCTACAGTAGCAGTCCCTAAATCTGGAATTTGGTGTGGTGGGATGATGCCATC 423
 QY 183 HisTyrGlyGlyProLeuArgLeuGlnPheTyrLeuIleThrMetSerAlaLeuMet 202
 Db 424 CACTGGAAGGCCCCCTTCGACTGCAGCAGCGCTATCTCATTATGATCAGTGCCTCATG 483
 QY 203 AlaLeuValPheIleLysTyrLeuProGluTyrThrValTyrPheValLeuPheValIle 222
 Db 484 GCCTCGTATTATCAAGTACCTCCCGAATGACCGCATGGCTCATCTTTGGGTGTGATT 543
 QY 223 SerValTyrAspLeuValAlaValLeuThrProLysGlyProLeuArgTyrLeuValGlu 242
 Db 544 TCAGTATATGATTTGGTGGCTGTTTATGTCNCAAGGCCCATCTTGGTATGCTGTTGAA 603
 QY 243 ThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSerGlyValIle 262
 Db 604 ACAGCTCAGGAAGAAATGAGACTCTCTTCCAGCTCTTATCTATCTCTCAACANTGGTG 663
 QY 263 TyrProTyrValLeuValThrAlaValGlu 272
 Db 664 TGG-----TTGGTGAATATGGCTGAA 684

RESULT 12
 AI925372
 LOCUS
 DEFINITION
 wn53d06.x1 NCT CGAP Lu19 Homo sapiens cDNA clone IMAGE:249163 3'
 similar to SW:FSN2_HUMAN P49810 PRESENILIN 2 ; mRNA sequence.
 ACCESSION
 AI925372
 VERSION
 AI925372.1 GI:5661336
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 703)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL
 Unpublished
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 459.
 Location/Qualifiers
 1. .703
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:249163"
 /tissue_type="squamous cell carcinoma, poorly
 differentiated (4 pooled tumors, including primary and
 metastatic)"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI-CGAP Lu19"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 pooled lung tumor tissue, and was then primed with a Not I
 - oligo(dT) primer. Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and

FEATURES
 source


```

603 GAGCCCATATTCCTCGCCCTGATATATATCATCTATCTGCCATGCTGTGGACGGTGGCATGGCG 544
QY
269 Thr-----AlaValGluAsnThrThrAspProArg-----GluPro 280
Db
543 AAGCTGAGACCCCTCTCTCAGGTGCTCCCTCCACTCCCTACACCCGGAGATGGAGAA 484
QY
281 ThrSerSerAspSerAsnThrSerThrAlaPheProGlyGluAlaSerCysSerSerGlu 300
Db
483 GACTCCTCATGACAGTTT-----GGGAG-----460
QY
301 ThrProLysArgProLysValLysArgIleProGlnLysValGlnIleGluSerAsnThr 320
Db
459 ---CCTTCATACCCCGAAGTCTTGGACCTCCC-----430
QY
321 ThrAlaSerThrThrGlnAsnSerGlyValArgValGluArgGluLeuAlaAlaGluArg 340
Db
429 -----TTGACTGCTACCCAGGGAGAGCTGGAG-----400
QY
341 ProThrValGlnAspAlaAsnPheHisArgHisGluGluGluArgGlyValLysLeu 360
Db
399 -----GAGAGGAGGAAAGGGCGTGAAGCTT 373
QY
361 GlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAla-----SerSerTyr 378
Db
372 GGCCTCGGGGACTTCATCTTCTACAGTGTGCTGTGGCAGGCGGTGCCACGGCGACGG 313
QY
379 PheAspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGlyLeuCysPheThr 398
Db
312 GG-GACTGGATACACGCTGGCTGTCTGTGGCCATCCTCATTTGGCTTGTGTGACC 254
QY
399 LeuValLeuLeuAlaValPheLysArgAlaLeuProAlaLeuGln-----413
Db
253 CTCCTGCTGTGTGTGTTCAGAAAGCGCTGCGCGCCCTCCCATCTCCATCACGTTTC 194
QY
414 -----PheProPheSerProAspSer 420
Db
193 GGGCTCATCTTTTACTTCTCCACGGACAAAC 164

RESULT 14
LOCUS
DEFINITION
  A1831581 680 bp mRNA linear EST 21-DEC-1999
  similar to SW:PSN2_HUMAN P49810 PRESENILIN 2, mRNA sequence.
ACCESSION
  A1831581
VERSION
  A1831581.1 GI:5452252
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 680)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cga@bbs-r@mail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 1711 std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 456.
  Location/Qualifiers
    1..680
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"

FEATURES
  source

```

```

/clone="IMAGE:2405191"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev stage="adult"
/lab_host="NCI CGAP Lu19"
/clone_lib="NCI CGAP Lu19"
/notes="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo (dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 137 a 191 c 173 g 178 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.24e-57 Length: 680
Score: 681.50 Matches: 131
Percent Similarity: 78.28% Conservative: 42
Best Local Similarity: 59.28% Mismatches: 47
Query Match: 28.66% Indels: 1
DB: Gaps: 1
US-09-043-944-1 (1-461) x A1831581 (1-680)
QY 43 LeuLysTyrGlyAlaSerHisValIleHisLeuPheValProValSerLeuCysMetAla 62
Db 9 CTCAAATACTGATCGAAGCAGCTGATCATGCTGTTGTGCTGTCTCTCTGTGCATGATC 68
QY 63 LeuValValPheThrMetAsnThrIleThrPheTyrSerGlnAsnAsnGlyArgHisLeu 82
Db 69 GTGGTGTAGCCACCATCAAGTCTGTGGCTTCTACACAGAGAAAGATGGACAG---CTC 125
QY 83 LeuSerHisProPheValArgGluThrAspSerIleValGluLysGlyLeuMetSerLeu 102
Db 126 ATCTACACGGCCATTCACTGAGGACACACCCCTCGGTGGCCAGCGCCTCTCACTCCGTG 185
QY 103 GlyAsnAlaLeuValMetLeuCysValValValLeuMetThrValLeuLeuIleValPhe 122
Db 186 CTGAACACCCCTCATCATGATCAGCGTATCTGTTGTTATGACCATCTTCTTGTGTGCTC 245
QY 123 TyrLysTyrLysPheTyrLysLeuIleHisGlyTrpLeuIleValSerSerPheLeuLeu 142
Db 246 TACAAGTACCCTGCTACAGTTTCATCATGGCTGGTTGATCATGCTCTTCACTGATGCTG 305
QY 143 LeuPheLeuPheThrThrIleTyrValGlnGluValLeuLysSerPheAspValSerPro 162
Db 306 CTGTTCTCTTTCACCTATATCTACTTGGGGAAGTGTCTAAGACCTACAATGTGGCCATG 365
QY 163 SerAlaLeuLeuValLeuPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMetCysIle 182
Db 366 GACTACCCACCCCTCTGCTGACTGCTGGAACCTCGGGCAGTGGGATGTTGTGTGATC 425
QY 183 HisTrpLysGlyProLeuArgLeuGlnPheTyrLeuIleThrMetSerAlaLeuMet 202
Db 426 CACTGGAAGGGCCCTCTGCTGCTGAGCAGGCTACCTCATCATGATCAGTGGCTCATG 485
QY 203 AlaLeuValPheIleTyrLeuProGluThrPheValTrpPheValLeuPheValIle 222
Db 486 GCCCTAGTGTTCATCAAGTACTCCAGAGTGGTCCGGGTGGTGGTCTACTGNGGCCATC 545
QY 223 SerValTrpAspLeuValAlaValLeuThrProLysGlyProLeuArgTyrLeuValGlu 242
Db 546 TCTGTGTATGATCTCGTGGCTGTGCTGTGTACCATAGGGCTCTGAGAAATGCTGGTAG 605
QY 243 ThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSerSerGlyValIle 262
Db 606 ACTGCCAGGAGAGATGAGCCCATATTCCCTGCTGATATATCTCATCTGACATGTTGTG 665

```

```

QY      263 Tyr 263
Db      666 TGG 668

RESULT 15
LOCUS   CD358381
DEFINITION AGENCOURT_14284651 NIH_MGC_180 Homo sapiens cDNA clone
IMAGE:30389377 5', mRNA sequence.
ACCESSION CD358381
VERSION   CD358381.1 GI:31129816
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 881)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
           Office of Cancer Genomics
           National Cancer Institute / NIH
           Bldg. 31 Rm10A07 Bethesda, MD 20892
           Email: cgabbs-xemail.nih.gov
           Tissue Procurement: Dr. Michael Brownstein
           cDNA Library Preparation: Invitrogen Corp
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LIML at:
           http://image.llnl.gov
           Plate: NDAM462 row: m column: 02
           High quality sequence stop: 685.

FEATURES             source
     1..881
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clones="IMAGE:30389377"
     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
     /clone_lib="NIH_MGC 180"
     /note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;
     Site 2: EcoRV (destroyed); Library is oligo-dT primed and
     directionally cloned (EcoRV site is destroyed upon cloning
     ). Average insert size 1.68 kb. Library was constructed by
     (Invitrogen). Note: this is a NIH_MGC Library."

BASE COUNT  165 a 254 c 252 g 210 t
ORIGIN

```

Search completed: December 11, 2003, 08:21:17
Job time : 2296 secs

```

182 TTCACTATATCTACCTTGGGGAAGTGTCTCAAGACCTACAAATGTGGCCACTGACTACCCC 241
166 LeuValLeuPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMetCysIleHisTrpLys 185
242 ACCTCTTGTGACTGTCTGTGGAACCTTCGGGGCAGTGGGCATGGTGTGCATCCACTGGAAG 301
186 GlyProLeuArgLeuGlnGlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuVal 205
302 GGCCCTCTGTGCTGTCAGCAGGCTACCTCATCATGATCATGCTGCTCATGGCCCTAGTG 361
206 PheIleLysTyrLeuProGluTyrTrpPheValLeuPheValIleSerValTrp 225
362 TTTCATCAAGTACCTCCAGAGTGGTCCGCGTGGTTCATCTCTGGCGCCATCTCTGTGTAT 421
226 AspLeuValAlaValLeuThrProLysGlyProLeuArgTyrLeuValGluThrAlaGln 245
422 GATCTCTGGTGTGTGTGTCTCCAAAGGGCTCTGAGAATGCTGTGTAAGAACTGCCCAG 481
246 GluArgAsnGluProIlePheProAlaLeuIleTyrSerSerGlyValIleTyrProTyr 265
482 GAGAGAAATGAGCCCATATTCCTGCTGCTGCTGATATCTCATCTGCTGCTGGTGGAGGTT 541
266 ValLeuValThrAlaValGluAsnThrThrAsp-----ProArgGluProThr 281
542 GGATGGCGAAGCTGGACCCCTCTCTCAGGTGCCCTCCAGCTCCCTACGACCCGGAG 601
282 SerSerAspSerAsnThrSerThrAlaPheProGlyGluAlaSerCysSerSerGluThr 301
602 ATGGAAGACTCCTATGACAGTTT-----GGGGAG----- 631
302 ProLysArgProLysValLysArgIleProGlnLysValGlnIleGluSerAsnThrThr 321
632 CCTTCATACCCCGAAGTCTTTGAGCCTCCC----- 661
322 AlaSerThrThrGlnAsnSerSerGlyValArgValGluArgGluLeuAlaAlaGluArgPro 341
662 -----TTGACTGGCTACCCAGGGAGGAGCTGGAG----- 691
342 ThrValGlnAspAlaAsnPheHisArgHisGluGluGluArgGlyValLysLeuGly 361
692 -----GAGAGAGGAGGAAGGGCGTGAAGCTTGGC 721
362 LeuGlyAspPheIlePheTyrSerValLeuLeuGly 373
722 CTCGGGAGACTTCATCTTCTACACTGTGATAGGGGCG 757

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2003, 06:24:06 ; Search time 339 Seconds
(without alignments)
3670.917 Million cell updates/sec

Title: US-09-043-944-1

Perfect score: 2378

Sequence: 1 MFSTRQGGGADAEHTV.....YYINSLFLPLCIINFSIIS 461

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/US09043944/runat_11122003_062359_3663/app.query.fasta_1.647
-DB=N_Geneseq 19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09043944 @CGN 1.1.490 @runat_11122003_062359_3663 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2378	100.0	1500	18	Caenorhabditis ele
2	2052.5	86.3	4137	24	C elegans sel-12 g
3	994.5	41.8	2765	18	Human S182 gene, p
4	993.5	41.8	1488	18	Partial AD3 sequen
5	993.5	41.8	1703	19	PS1/429 protein co
6	993.5	41.8	2764	19	PS1/467 protein co
7	993.5	41.8	2764	24	Presenilin coding
8	993.5	41.8	2765	17	Presenilin-1-1 wil
9	993.5	41.8	2765	19	Human presenilin-1
10	993.5	41.8	2765	22	Human presenilin-1
11	993.5	41.8	2765	22	Nucleotide sequenc
12	991	41.7	1750	18	Human early onset
13	991	41.7	1762	18	Human early onset
14	989.5	41.6	3056	24	Human cDNA differe
15	989.5	41.6	3086	17	Presenilin-1-2, al
16	989.5	41.6	3086	19	Human presenilin-1
17	989	41.6	1964	17	Murine presenilin-
18	989	41.6	1964	19	Mouse presenilin-1
19	989	41.6	2681	18	Tumour suppressor
20	988.5	41.6	1895	17	Presenilin homolog
21	988.5	41.6	1911	18	Human S182 gene as
22	988.5	41.6	1914	18	Presenilin-1 VRSQ
23	988.5	41.6	2048	23	Drosophila melanog
24	986.5	41.5	2764	19	Homo sapiens PS-1
25	985.5	41.4	1404	20	Mouse presenilin-1
26	983	41.3	1750	19	Human presenilin I
27	979.5	41.2	2765	18	Human mutant S182
28	979	41.2	1392	19	Human presenilin I
29	979	41.2	1404	20	Human presenilin I
30	979	41.2	1404	22	Human presenilin I
31	970.5	40.8	2276	18	Full AB4/AD3LP seq
32	969	40.7	1404	24	Human mutant prese
33	967	40.7	2229	17	Human presenilin-2
34	967	40.7	2229	19	Human presenilin-2
35	967	40.7	2236	18	Human AD4 protein
36	964.5	40.6	1705	25	Human signalling p
37	959	40.3	2144	21	Human presenilin-2
38	959	40.3	2144	25	Human PS2 cDNA seq
39	959	40.3	2236	19	Human presenilin I
40	959	40.3	2236	22	Human presenilin-2
41	959	40.3	2236	22	Nucleotide sequenc
42	958.5	40.3	2527	22	Human EST-derived
43	950.5	40.0	1983	21	Human presenilin-2
44	949	39.9	1347	24	Human mutant prese
45	947	39.8	1347	24	Human mutant prese

ALIGNMENTS

RESULT 1

AAT60306

ID AAT60306 standard; DNA; 1500 BP.

XX AAT60306;

AC AAT60306;

XX 11-JUL-1997 (first entry)

DT 11-JUL-1997 (first entry)

XX Caenorhabditis elegans SEL-12 gene.

DE Caenorhabditis elegans SEL-12 gene.

XX SEL-12; presenilin; transgenic animal; Alzheimer's disease;

KW animal model; therapy; diagnosis; ss.

XX Caenorhabditis elegans.

OS Caenorhabditis elegans.

XX Key Location/Qualifiers

FH CDS 22..1407


```

FT      /*tag= a
XX      WO9711956-A1.
XX      03-APR-1997.
XX      27-SEP-1996; 96WO-US15727.
XX      27-SEP-1995; 95US-0004387.
XX      (UYCO ) UNIV COLUMBIA NEW YORK.
XX      Greenwald I, Levitan D;
XX      WPI; 1997-212841/19.
XX      P-PSDB; AAW14006.
XX
XX      New isolated SEL-12 presenilin gene - used to develop products for
XX      the diagnosis and prophylactic or therapeutic treatment of disorders
XX      such as Alzheimer's disease
XX
XX      Claim 11; Page 75-76; 99pp; English.
XX
XX      The sel-12 gene (AAW60306) of Caenorhabditis elegans (CE) encodes a
XX      polypeptide (AAW14006) that displays about 50% amino acid sequence
XX      identity to human presenilin PS1 and PS2 proteins associated with
XX      Alzheimer's disease (AD). Like presenilins, sel-12 is widely
XX      expressed in neural and non-neural cells. The gene was identified
XX      by screening for suppressors of the 'Multivulva' phenotype of CE
XX      caused by an allele of lin-12 that causes constitutive LIN-12
XX      activation. 2 Recessive suppressors, arl31 and arl33, proved to be
XX      alleles of the new gene (sel means suppressor and/or enhancer of
XX      lin-12). The gene can be used to produce wild-type or mutated (see
XX      also AAW14007) SEL-12 in host cells, in methods for screening for
XX      extragenic suppressors or enhancers of a SEL-12 allele, and in the
XX      development of transgenic animals, esp. transgenic CE, useful in
XX      identifying cpds. that ameliorate AD.
XX
XX      SQ      Sequence 1500 BP; 403 A; 312 C; 314 G; 471 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.:      3,946-230      Length:      1500
XX      Score:          2378.00      Matches:      461
XX      Percent Similarity: 100.00%      Conservative: 0
XX      Best Local Similarity: 100.00%      Mismatches: 0
XX      Query Match:      100.00%      Indels:      0
XX      DB:              18      Gaps:      0
XX
XX      US-09-043-944-1 (1-461) x AAW60306 (1-1500)
XX
XX      QY      1 MetProSerThrArgGlnGlnGluGlyGlyAlaAspAlaGluThrHisThrVal 20
XX      DB      22 ATGCTTCCACAGAGAGACACAGAGGCGGAGGTGCAGATCGGAAACACATACCGTT 81
XX
XX      QY      21 TyrGlyThrAsnLeuIleThrAsnArgAsnSerGlnGluAspGluAsnValGluGlu 40
XX      DB      82 TACGGTACAAATCTCATACAAATCGAATAGCCAAAGAGAGAGAAATGTTGTGGAAGAA 141
XX
XX      QY      41 AlaGluLeuLysTyrGlyAlaSerHisValIleHisLeuPheValProValSerLeuCys 60
XX      DB      142 GCGGAGCTGAAATACGAGACATCTCACGTTATCACTATTGTCGCGGTGCTACTATGC 201
XX
XX      QY      61 MetAlaLeuValValPheThrMetAsnThrIleThrPheTyrSerGlnAsnAsnGlyArg 80
XX      DB      202 ATGGCTCTGTTGTTTACGATGACACAGATTACGTTTATAGTCAAAACAATGGAAGG 261
XX
XX      QY      81 HisLeuLeuSerHisProPheValArgGlnThrAspSerIleValGluLysGlyLeuMet 100
XX      DB      262 CATTTACTATCATCCCTTTTGTCCGGAACACAGACATATCGTTGGAAGGGATGATG 321
XX
XX      QY      101 SerLeuGlyAsnAlaLeuValMetLeuValMetLeuValValLeuMetThrValLeuLeu 120
XX      DB      322 TCACITGGAAATGCTCTGTCATGTTGTGCGTGGTTCGTCTGTGATGACAGTTCTGCTGATT 381

```

```

QY      121 ValPheTyrLysTyrLysPheTyrLysLeuIleHisGlyTyrLeuIleValSerSerPhe 140
DB      382 GTTTCATATAAATACAAGTTTTATTAAGCTTATTCATGAATGCTTATTTGTCAGCAGTTT 441
QY      141 LeuLeuLeuPheLeuPheThrThrIleTyrValGlnGluValLeuLysSerPheAspVal 160
DB      442 CTTCTCTTTTCTTATTCACATCAATCTATGTGCAAGAAGTTCTGAAAGTTTCGATGTG 501
QY      161 SerProSerAlaLeuLeuValLeuPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMet 180
DB      502 TCTCCAGCGCATTTGGTTTGTGGACTGGGTAACATATGGAGTTCTCGGAATGAG 561
QY      181 CysIleHisTyrLysGlyProLeuArgLeuGlnGlnPheTyrLeuIleThrMetSerAla 200
DB      562 TGTATACATGGAAAGGTCCATTGGCTCTGCACAGTTCTACCTTATTACAATGTCTGCA 621
QY      201 LeuMetAlaLeuValPheIleLysTyrLeuProGluTyrThrValTyrPheValLeuPhe 220
DB      622 CTAAATGGCTCTGGTCTCTTATCAAGTACCTACAGAAATGGACTGTGTGGTTTGTGCTGTT 681
QY      221 ValIleSerValTyrAspLeuValAlaValLeuThrProLysGlyProLeuArgTyrLeu 240
DB      682 GTTATCTCGTTTGGATCTGGTTCCGCTGCTCACACCAAAAGGACCATTTGAGATATTG 741
QY      241 ValGluThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSerSerGly 260
DB      742 GTGGAACTGTCACAGAGAGAAACGAGCCAAATTTTCCGCGCTGATTTATTTCGTCTGGA 801
QY      261 ValIleTyrProTyrValLeuValThrAlaValGluAsnThrThrAspProArgGluPro 280
DB      802 GTCATCTATCTTCCTAGCTTCTTGTCTGAGTTGAAACACAGACAGACCCCGTGAACCG 861
QY      281 ThrSerSerAspSerAsnThrSerThrAlaPheProGlyGluAlaSerCysSerSerGlu 300
DB      862 AGTCGTGAGACTCAAAATCTTCTACAGCTTTCTCTGGAGAGCGAGTTGTTTCATCTGAA 921
QY      301 ThrProLysArgProLysValLysArgIleProGlnLysValGlnIleGluSerAsnThr 320
DB      922 AGCCCAAAACGCGCAAAAGTGAACGAAATTCCTCAAAAGTCAAAATCGAATCGAATACT 981
QY      321 ThrAlaSerThrThrGlnAsnSerGlyValArgValGluArgGluLeuAlaAlaGluArg 340
DB      982 ACAGCTTCAACGACACACAAACTCTGGAGTAAAGGTGGAAACGCGAGCTAGCTGTCAGAGA 1041
QY      341 ProThrValGlnAspAlaAsnPheHisArgHisGluGluGluArgGlyValLysLeu 360
DB      1042 CCAACTGTACAGACGCGCCAAATTTTACAGGACACGAAAGAGAGAGAGAGGTGTGAACCTT 1101
QY      361 GlyLeuGlyAspPheIlePheTyrSerValLeuLeuLeuGlyLysAlaSerSerTyrPheAsp 380
DB      1102 GGTCTGGGCGACTTCATTTTCTACTCTGTTCTCTCTCGGCAAGGCTTCATCGTACTTTGAC 1161
QY      381 TrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGlyLeuCysPheThrLeuVal 400
DB      1162 TGGAAACACGACTATCGCTTGTATGTGGCCATTTCTTTCGGTCTCTGCTTCACTTGTCTC 1221
QY      401 LeuLeuAlaValPheLysArgAlaLeuProAlaLeuGlnPheProPheSerProAspSer 420
DB      1222 CTGCTCGCGCTCTTCAAGAGACACTCCCGGCTCTGCAATTTTCCATTTTCTCCGAGCTCA 1281
QY      421 PhePheThrPheValProAlaGlySerSerProHisLeuLeuHisLysSerLeuLysSer 440
DB      1282 TTTTTCATTTTGTACCGCTGATCATCATCCCCCATTTGTTTACACAAGTCTCTCAAAAGT 1341
QY      441 ValTyrTyrIleAsnSerLeuPheLeuProPheLeuCysIleIleAsnPheSerIleIle 460
DB      1342 GTTATATATATTAATCTCTGTTTGTGCAATTTCTTTCATCATCATCACTTTTTCGATTATA 1401
QY      461 Ser 461
DB      1402 TCT 1404

```

RESULT 2

AAL47322 standard; DNA; 4137 BP.
 ID AAL47322; AC AAL47322; DT 02-SEP-2002 (first entry)
 XX C elegans sel-12 gene promoter and regulatory regions.
 XX Sel-12; presenilin; neuronal disorder; familial Alzheimer's disease;
 XX amyloid precursor protein; APP; ds.
 XX Caenorhabditis elegans.
 XX US6376239-B1.
 XX 23-APR-2002.
 XX 04-APR-1997; 97US-0832867.
 XX 04-APR-1997; 97US-0832867.
 XX (ELEG-) ELEGNE GMBH.
 XX Baumeister R;
 XX WPI; 2002-478281/51.
 XX Isolated DNA molecule comprising promoter of the sel-12 gene from
 PT Caenorhabditis elegans operably linked to heterologous gene, directs
 PT expression in neural cells and is useful to develop drugs to treat
 PT neuronal disorders -
 XX Claim 1; Fig 3; 78pp; English.
 PS
 CC The present invention relates to DNA molecules comprising the promoter of
 CC the sel-12 gene from Caenorhabditis elegans operably linked to a
 CC heterologous DNA sequence encoding a protein of interest. The sequence
 CC can be used to develop drugs for the treatment, prevention or delay of a
 CC neuronal disorder. In particular, the neuronal disorder may be familial
 CC Alzheimer's disease. The present sequence is the C. elegans sel-12
 CC promoter.
 XX
 SQ Sequence 4137 BP; 1252 A; 770 C; 703 G; 1412 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.4e-196 Length: 4137
 Score: 2052.50 Matches: 459
 Percent Similarity: 58.82% Conservative: 1
 Best Local Similarity: 58.70% Mismatches: 1
 Query Match: 86.31% Indels: 322
 DB: 24 Gaps: 7
 US-09-043-944-1 (1-461) x AAL47322 (1-4137)
 QY 1 MetProSerThrArgArgGlnGlnGluGlyGlyAlaAspAlaGluThrHis----- 18
 DB 960 ATGCTTCCACAGGACACACAGGAGGCGGAGGTGCAGATCGGAAACACATGTAAGT 1019
 QY 18 ----- 18
 DB 1020 TATTTAGACATTTTATTTTCTCAAGAACTAAATTTGTTAAATTTGCTACAATGCTTT 1079
 QY 19 -----ThrValTyrGlyThrAsnLeuIleThrAsnArgAsnSerGlnGluAspGluAsnVa 37
 DB 1080 TCACACGTTTACGGTACAAATCTGTATACAAATCGGAATAGCAGAGACGAAATGT 1139
 QY 37 lValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIleHisLeuPheValProVa 57
 DB 1140 TGTGGAGAGCGGAGCTGAAATACGAGGATCTCAGTTTATTCATCTATTGTGCGGT 1199
 QY 57 lSerLeuCysMetAlaLeuValPheThrMetAsnThrIleThrPheTyrSerGlnAs 77

Db 1200 GTCACTATGCAATGGCTCTGGTGTGTTTACATGACACACGATTACGTTTATAGTCAAAA 1259
 QY 77 nAsnGlyArgHisLeu----- 82
 Db 1260 CAATGGAAGGCATTTGTAAAGTTTCTAAAGAAATTCATGATTAAAAATATTGATTTGTT 1319
 QY 83 -----LeuSer----- 84
 Db 1320 TTATCAATTTGCACTCTGGCAATCGCACTCTTTGTGTCAGTGCAGAAATAAATTTTGGTCAGT 1379
 QY 84 ----- 84
 Db 1380 GCAAAATAATTTTGGTCAGTGCATTTGATATTATGTCAGTGCATTTGCAAGTCGAGCT 1439
 QY 84 ----- 84
 Db 1440 TTAATAATTTTCGTGGTGTAAATTTTACTCAATTTCTATCAATATTCTTTTGGAAAAAA 1499
 QY 84 ----- 84
 Db 1500 GTTGAAGATTTACTCTGGAAATTTTCGAAATAAATCTGTAATGGAATAATCAATCAACACA 1559
 QY 85 -----His-ProPheValArgGluThrAspSerIleValG 96
 Db 1560 AACTTTGAATATTATTTTCAGACTATACATCTCCTTTTGTCCGGGAAACACACAGTAGTATCGTTG 1619
 QY 96 lLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysValValValLeuMetT 116
 Db 1620 AGAAGGATGATGTCACCTCTGGAATGCTCTGCAATGTTGTGCGTGTCTTCTGATGA 1679
 QY 116 hrValLeuLeuValPheTyrLysPheTyrLysLeuLeuHisGlyTrpLeuI 136
 Db 1680 CAGTTCTGCTGATTTTCTATAAATACAAAGTTTATAAGCTTATTTCATGATGGCTTA 1739
 QY 136 lValSerSerPheLeuLeuPheLeuPheThrThrIleTyrValGln----- 152
 Db 1740 TTGTCAGCAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1799
 QY 153 -----GluValLeuLysSerPheAspValS 161
 Db 1800 TATTACTATTCTCATAAAAATATCAATGTTGTCAGAGAAGTTCTGAAAGTTTCTCGAATGATGT 1859
 QY 161 erProSerAlaLeuLeuValLeuPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMetC 181
 Db 1860 CTCCACGCGCATATTTGTTTGTGACTGGGTAACTATGGAGTTCTCGAATGATGT 1919
 QY 181 ysIleHisTrpLysGlyProLeuArgLeuGlnGlnPheTyrLeuIleThrMetSerAla 201
 Db 1920 GTATACATTTGGAAGGTCCATTTGGTCTGCAACAGTTCTACCTTATTACAAATGTCGAC 1979
 QY 201 euMetAlaLeuValPheIleLysTyrLeuProGluTrpThrValTrpPheValLeuPheV 221
 Db 1980 TAAATGGCTCTGCTCTTTATCAAGTACCTACCAGAATGGACTGTGTGTTGTGCTGTTG 2039
 QY 221 alIleSerValTrpAspLeuValAlaValLeuThrProLysGlyProLeuArgTyrLeuV 241
 Db 2040 TTATCTCGGTTTGGATCTGTTTCCGCTCTCACCAAAAGGACCAATTGAGATATTG 2099
 QY 241 alGluThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSer----- 258
 Db 2100 TGGAACTGACAGAGAGAAACGAGCCAAATTTTCCGCGCTGATTTTTCGTGTAAGT 2159
 QY 259 -----Ser-Gly 260
 Db 2160 TTCCTAATTTTGAATTAATATTATTCATGACGTTTCAAATTTCTAAACATTTTTCAGTCTGA 2219
 QY 261 ValIleTyrProTyrValLeuValThrAlaValGluAsnThrThrAspProAlaGluPro 280
 Db 2220 GTCATCTATCCTACGTTCTTGTATTGTCAGTTGAAACACAGACAGACCCCGGTGAACCG 2279
 QY 281 ThrSerSerAspSerAsn----- 286

Score:	993.50	Matches:	216
Percent Similarity:	65.84%	Conservative:	77
Best Local Similarity:	48.54%	Mismatches:	104
Query Match:	41.78%	Indels:	48
DB:	18	Gaps:	9

US-09-043-944-1 (1-461) x AAT87402 (1-1488)

Qy	32	GlnGluAspGluAsnValValIgluclAlaGluLeuLysTyrGlyAlaSerHisVallle	51
Db	11	GAGCAAGATCAGAAAGAAGATGAGAGCTGCATTTGAAATATCGCCACGAAGCATGTGATC	70
Qy	52	HisLeuPheValProValSerLeuCysMetAlaLeuValPheThrMetAsnThrIle	71
Db	71	ARGCTCTTTGTCCCTTGACTCTTCGATGTTGGTGTGGTACCATTAAAGTCAGTC	130
Qy	72	ThrPheTyrSerGlnAsnGlyArgHisLeuLeuSerHisPropheValArgGluThr	91
Db	131	AGCTTTTATACCCGAAGATGGGCAG---CTAATCTATACCCCAITTCACAGAATAAC	187
Qy	92	AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal	111
Db	188	GAGACTGTGGCCAGAGAGCCCTGCATCAATTCGATGCTGCCATCATGATCAGTGC	247
Qy	112	ValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuIle	131
Db	248	ATTGTTGTCA TGACTATCCTCTCGTGTCTGTATAAATACAGGTGCTATTAAGGTCATC	307
Qy	132	HisGlyTrpLeuIleValSerPheIleLeuLeuPheLeuPheThrIleTyrVal	151
Db	308	CATGCGCTGGCTTATTATATCATCTCTATGTGTGTGTTCTTTTTTCATTCA TTACTTG	367
Qy	152	GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu	171
Db	368	GGGGAAGTGT TAAAACCTATTAACGTTCGCTGGAGTACATTACTGTTGCACCTCGATC	427
Qy	172	GlyAsnTyrGlyValLeuGlyMetCyslIleHisTrpLysGlyProLeuArgLeuGln	191
Db	428	TGGAATTTTGGTGTGGTGGAAATGATTTCCTACTCTGGAAGAGTCCACTTCGACTCCAG	487
Qy	192	GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro	211
Db	488	CAGGCATATCTCATATGATGATAGTGCCTCATGGCCCTGTGTGTTTATCAAGTACCTCC	547
Qy	212	GluTrpThrValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeu	231
Db	548	GAATGGACTCGGTGGCTCATCTTGGCTGTGATTTTCAGTATATGATTTAGTGGCTGTTTG	607
Qy	232	ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle	251
Db	608	TGTCGGAAGGTCATCTCGTATGCTGGTGTGAACAGCTCAGGAGAGAAATGAACGC	667
Qy	252	PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal	271
Db	668	TTTCCAGCTCTCATTTTACTCTCCAACATGGGTGG-----TTGGTGAATATGGCA	718
Qy	272	GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe	291
Db	719	GAAGGAGACCCGGAAGCTCAAAGAGAGATGTCCAAAAATTC-----	760
Qy	292	ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro	311
Db	761	-----AAGTATATATGCAGAAAGCACAGNAAGG-----	787
Qy	312	GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg	331
Db	788	-----GAGTCAACAAGACACTGTTGCAGAGAAATGATGTCGGGTTCAGT	832
Qy	332	ValGlu-----	339
Db	833	GAGGATGGGAAGCCGAGGAGGACACTCATCTAGGCGCTCATGCTCTACACTTGAGTCA	892
Qy	340	ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluGluArg	356

DT 02-SEP-2002 (first entry)
 XX Presenilin coding sequence.
 DE
 XX Sel-12; presenilin; neuronal disorder; familial Alzheimer's disease;
 KW amyloid precursor protein; APP; ds.
 XX
 XX Unidentified.
 XX
 XX Key Location/Qualifiers
 FT CDS 249..1652
 FT /*tag= a
 FT /product= "presenilin"
 XX
 XX US6376239-B1.
 XX
 XX 23-APR-2002.
 XX
 XX 04-APR-1997; 97US-0832867.
 XX
 XX 04-APR-1997; 97US-0832867.
 XX
 XX (ELEG-) ELEGNE GMBH.
 XX
 XX Baumeister R;
 XX
 XX WPI; 2002-478281/51.
 XX
 XX P-PSDB; NAO18049.
 XX
 XX Isolated DNA molecule comprising promoter of the sel-12 gene from
 PT Caenorhabditis elegans operably linked to heterologous gene, directs
 PT expression in neural cells and is useful to develop drugs to treat
 PT neuronal disorders -
 XX
 XX Claim 2; Fig 4; 78pp; English.
 XX
 XX The present invention relates to DNA molecules comprising the promoter of
 CC the sel-12 gene from Caenorhabditis elegans operably linked to a
 CC heterologous DNA sequence encoding a protein of interest. The sequence
 CC can be used to develop drugs for the treatment, prevention or delay of a
 CC neuronal disorder. In particular, the neuronal disorder may be familial
 CC Alzheimer's disease. The present sequence is a presenilin coding sequence
 CC described in the exemplification of the invention.
 XX
 XX Sequence 2764 BP; 715 A; 624 C; 652 G; 773 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 7,78e-90 Length: 2764
 Score: 993.50 Matches: 216
 Percent Similarity: 65.84% Conservative: 77
 Best Local Similarity: 48.54% Mismatches: 104
 Query Match: 41.78% Indels: 48
 DB: 24 Gaps: 9
 US-09-043-944-1 (1-461) x AAL47323 (1-2764)
 QY 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValile 51
 DB 438 GAGCAAGATGAGGAGAGATGAGGAGCTGACATTTGAAATATGGCCAGCAATGATC 497
 QY 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrile 71
 DB 498 ATGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
 QY 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuSerHisProPheValArgGluThr 91
 DB 558 AGCTTTTATACCCGGAAGATGGGAG---CTATCTATACCCCATCCATCAAGAATACC 614
 QY 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
 DB 615 GAGACTGTGGCCAGAGAGCCCTGCACTCAATTCGAATCTGCCATCATGATGATGTC 674
 QY 112 ValValLeuMetThrValLeuLeuLeuValPheTyrLysTyrLysPheTyrLysLeuile 131

Db 675 ATTGTTGTCATGACTATCTCTCTGTTGGTCTGTATAAATACAGGTGCTATAAGTCTATC 734
 QY 132 HisGlyTyrLeuLeuValSerSerPheLeuLeuLeuPheLeuPheThrThrIleTyrVal 151
 Db 735 CATGCTGGCTTATATATATCATCTCTATATGTTGCTGTTCTTTTTCATTCATTCATTCG 794
 QY 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuValLeuPheGlyLeu 171
 Db 795 GGGGAAGTGTAAACCTTAACCTTCTGCTGGACTACATCTGTTGCACTCTCTGATC 854
 QY 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisThrLysGlyProLeuArgLeuGln 191
 Db 855 TGGAAATTTGTTGTTGGTGGAAATGATTTCCATTCATCGAAAGGTCCACTTCGACTCCAG 914
 QY 192 GlnPheTyrLeuLeuThrMetSerAlaLeuMetAlaLeuValPheLeuLysTyrLeuPro 211
 Db 915 CAGGCATATCTCATTTATGATTAGTCCCTCATGCTGGCTGTTATCATCAATCTCTCT 974
 QY 212 GluTyrThrValTyrPheValLeuPheValIleSerValTyrAspLeuValAlaValLeu 231
 Db 975 GAATGACTGCTGCTCATCTTGGCTGTGATTTTCAGTATATGATTTAGTGGCTGTTTG 1034
 QY 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
 Db 1035 TGTCCGAAAGGTCCACTTCTGATGCTGTTGAAACAGCTCAGGAGAGAAATGAACGCTT 1094
 QY 252 PheProAlaLeuLeuTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
 Db 1095 TTTCCAGCTCTCATTTACTCTCAACATGTTGGTGG-----TTGGTGAATATGGCA 1145
 QY 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
 Db 1146 GAAAGAGAGCCCGAAGCTCAAGAGAGATGATCAAAAATTC----- 1187
 QY 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
 Db 1188 -----AAGTATAATGCAAGAACACAGAAAG----- 1214
 QY 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331
 Db 1215 -----GAGTCACAGACACTGTTCGAGAGATGATGATGCGGGTTCAGT 1259
 QY 332 ValGlu-----ArgGluLeuAlaGlu 339
 Db 1260 GAGGAATGGGAAGCCAGAGGAGCAGTCATCTAGGGCTCATCTGTACCTGTAGTCA 1319
 QY 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
 Db 1320 CGAGCTGCTGCTCCAGAACTTCCAGCAGTATCTCTCGCTGGTGAAGACCCAGAGAAAGG 1379
 QY 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
 Db 1380 GGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTCTGTTGGTGGTAAAGCTCA 1439
 QY 377 SerTyrPhe-----AspTyrAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
 Db 1440 GCACAGCCAGCTGGAGACTGGAAACACCACTAGCTGTTTCGTAGCCATTAATTTGGT 1499
 QY 395 LeuCysPheThrLeuValLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPh 414
 Db 1500 TTGTGCTTACATTTACTCTTGGCACTTTTCAAGAAAGCATTTGCCAGCTCTTCCAATC 1559
 QY 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLe 434
 Db 1560 TCCA---TCACCTTGGCTGTTTCTTACTTTCACAGATTAATCTGTACAGCCCTTT 1616
 QY 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh 451
 Db 1617 ATGGACCAATTAGCATTCATCAATTTATATCTAGCATATTTCCGGTGTAGAAATCCCATG 1676
 QY 451 eleuCysIlelele 455
 : : : : :

[illegible]

```

Db 1380 GGAGTAAACTGCTGATGGGAGATTCTCATTTCTACAGTCTTCTGGTGTGTAAGCCTCA 1439
Qy 377 SerTyrPhe-----AspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
Db 1440 GCAACAGCCAGTGGAGACTGGAACACACACATAGCCCTGTTTCGTAGCCATATTAATTTGGT 1499
Qy 395 LeuCysPheThrLeuValLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPh 414
Db 1500 TTGTGCTTACATATTAATTTACCTTCCCTTGGCCATTTCAAGAAACATGTCAGCTCTTCCAATC 1559
Qy 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLe 434
Db 1560 TCCA---TCACCTTTGGGCTGTTTCTTCTACTTTGCCACAGTATCTTGTACAGCCTTTT 1616
Qy 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh 451
Db 1617 ATGGACCAATTAGCATTCATCAATTTATATATCTAGCATATTTGCGGTTAGAAATCCCATG 1676
Qy 451 eLeuCysIleIle 455
Db 1677 GATGTTTCTTCTT 1689

RESULT 9
AAV04666
ID AAV04666 standard; cDNA; 2765 BP.
XX AC AAV04666;
XX DX 20-JUL-1998 (first entry)
XX DE Human presenilin-1 cDNA (hPS1-1).
XX KW Presenilin-1; PS1 gene; human; familial Alzheimer's disease; FAD;
XX KW cerebral haemorrhage; schizophrenia; depression; epilepsy;
XX KW mental retardation; diagnosis; therapy; transgenic animal; ss.
XX OS Homo sapiens.

Key Location/Qualifiers
FH 249..1652
CDS /*tag= a
FT exon 1..113
FT /*tag= b
FT /number= 1
FT 114..195
FT /*tag= c
FT /number= 3
FT 196..335
FT /*tag= d
FT /number= 4
FT 336..586
FT /*tag= e
FT /number= 5
FT 587..728
FT /*tag= f
FT /number= 6
FT 729..796
FT /*tag= g
FT /number= 7
FT 797..1017
FT /*tag= h
FT /number= 8
FT 1018..1116
FT /*tag= i
FT /number= 9
FT 1117..1203
FT /*tag= j
FT /number= 10
FT 1204..1377
FT /*tag= k
FT /number= 11
FT 1378..1496

```

```

FT exon 1
FT /number= 12
FT 1497..2765
FT /*tag= m
FT /number= 13
FT 324..335
FT /*tag= n
FT /note= "deletion of 12 nucleotides from 3' end of
FT exon 4 by alternative splicing, deletes
FT Val26-Gln-29 (not critical to PSI
FT function)"
FT 1018..1116
FT /*tag= o
FT /note= "absence of exon 9 through splicing
FT variation, results in Asp-257 changing
FT to Ala and fusion of Ala-257 to Thr-291"
FT 492
FT /*tag= p
FT /note= "G to C FAD mutation site (Val82Leu)"
FT 591
FT /*tag= q
FT /note= "T to C FAD mutation site (Tyr115His)"
FT 664
FT /*tag= r
FT /note= "T to C FAD mutation site (Met139Thr)"
FT 676
FT /*tag= s
FT /note= "T to C FAD mutation site (Ile143Thr)"
FT 684
FT /*tag= t
FT /note= "A to C FAD mutation site (Met146Leu)"
FT 736
FT /*tag= u
FT /note= "A to G FAD mutation site (His163Arg)"
FT 939
FT /*tag= v
FT /note= "G to A FAD mutation site (Ala231Thr)"
FT 985
FT /*tag= w
FT /note= "C to A FAD mutation site (Ala246Glu)"
FT 1027
FT /*tag= x
FT /note= "C to T FAD mutation site (Ala260Val)"
FT 1039
FT /*tag= y
FT /note= "C to T FAD mutation site (Pro264Leu)"
FT 1192
FT /*tag= z
FT /note= "C to T FAD mutation site (Ala285Val)"
FT 1104
FT /*tag= aa
FT /note= "C to G FAD mutation site (Leu286Val)"
FT 1399
FT /*tag= ab
FT /note= "G to C FAD mutation site (Gly384Ala)"
FT 1422
FT /*tag= ac
FT /note= "C to G FAD mutation site (Leu392Val)"
FT 1477
FT /*tag= ad
FT /note= "G to A FAD mutation site (Cys410Tyr)"
FT 1563
FT /*tag= ae
FT /note= "A to G FAD mutation site (Ile439Val)"
FT
XX WO9801549-A2.
XX 15-JAN-1998.
XX PD 04-JUL-1997; 97WO-CA00475.
XX PP 02-JAN-1997; 97US-0034590.
XX PR 05-JUL-1996; 96US-0021673.

```


Qy 414 eProPheSerProaspSerPhePheThrPheValProAlaGlySerSerProHisLeuLe 434
 Db 1560 TCCA---TCACCTTTGGGCTTTGTTTCTACTTTGGCCACAGATTATCTTGTACAGCCTTTT 1616
 Qy 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh 451
 Db 1617 ATGGACCAATAGCATTCATCAATTTATATCTAGCAATTTTGGGTTAGMATCCCATG 1676
 Qy 451 eLeuCysIleIle 455
 Db 1677 GATGTTTCTTCTT 1689

RESULT 11

AAH74993

ID AAH74993 standard; DNA; 2765 BP.

XX AAH74993;

DT 29-OCT-2001 (first entry)

XX Nucleotide sequence of human presenilin 1.

XX KIAA0253; presenilin; Alzheimer's disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FH 249..1652

FT CDS /*tag= a

FT FT /product= "presenilin 1"

PN W0200167109-Al.

XX 13-SEP-2001.

XX 09-MAR-2001; 2001WO-GB01057.

XX 10-MAR-2000; 2000GB-0005894.

XX (GLAX) GLAXO GROUP LTD.

XX Hale RS, Rowley A, Blackstock W;

XX WPI; 2001-522960/57.

XX P-PSDB; AAG63936.

PT Identifying a modulator of presenilin function by determining the
 PT ability of presenilin to bind to a KIAA0253 polypeptide in the presence
 PT and absence of a test compound, useful in the treatment or prophylaxis
 PT of Alzheimer's disease -

XX Disclosure; Page 42-44; 48pp; English.

XX The present sequence encodes human presenilin 1. KIAA0253 binds to
 CC presenilin. The specification describes a method of identifying a
 CC modulator of presenilin function or KIAA0253 function. The method
 CC comprises determining presenilin activity or KIAA0253 activity
 CC in the presence and absence of a test compound, where presenilin
 CC activity is determined by its ability to bind to KIAA0253. A modulator
 CC of presenilin or KIAA0253 polypeptide is useful in the manufacture of
 CC a medicament for the treatment or prophylaxis of Alzheimer's disease.
 CC The KIAA0253 polynucleotide and KIAA0253 polypeptide are useful in the
 CC treatment, prophylaxis or diagnosis of Alzheimer's disease.

XX Sequence 2765 BP; 715 A; 624 C; 652 G; 773 T; 1 other;

Alignment Scores:

Pred. No.: 7,78e-90 Length: 2765
 Score: 993.50 Matches: 216
 Percent Similarity: 65.84% Conservative: 77
 Best Local Similarity: 48.54% Mismatches: 104
 Query Match: 41.78% Indels: 48

Db: 22 Gaps: 9
 US-09-043-944-1 (1-461) x AAH74993 (1-2765)
 Qy 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIle 51
 Db 438 GAGCAAGATGAGGAGGAAGATGAGAGCTGCATTGAAATATGGCGCCAAAGCATGTGATC 497
 Qy 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrIle 71
 Db 498 ATGCTCTTTGTCCTGTGACTCTCTGCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 557
 Qy 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
 Db 558 AGCTTTTATACCCGGAAGATGGGCAG---CTAATCTATACCCCATTCACAGAAGATACC 614
 Qy 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
 Db 615 GAGACTGTGGCCAGAGAGCCCTGCACCTCAATCTGAATGCTGCTCATCATGATCATGTGC 674
 Qy 112 ValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuIle 131
 Db 675 ATTGTGTGATGACTATCTCTCTGGTGGTCTGTATATAATACAGGTGCTATAAGTGCATC 734
 Qy 132 HisGlyTyrLeuIleValSerSerPheLeuLeuLeuPheLeuPheThrIleTyrVal 151
 Db 735 CATGCTGGCTTATTATATCATCTCTATTGTGTGCTGTTCTTTTTCATTCACTTACTTG 794
 Qy 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValPheGlyLeu 171
 Db 795 GGGGAAGTGTTTAAACCTATACGTTCTGTGACTACATTACTGTGTGCATCTCTGTATC 854
 Qy 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTyrLysGlyProLeuArgLeuGln 191
 Db 855 TGGAAATTTGTGTGGTGGGAATGATTTCCTTCACTGGAAGGTCCACTTCGATCCAG 914
 Qy 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
 Db 915 CAGGCATATCTCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 974
 Qy 212 GluTyrThrValTyrPheValLeuPheValIleSerValTyrPheValValValVal 231
 Db 975 GAAIAGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1034
 Qy 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
 Db 1035 TGTCGGAAGGTCCACTTCGTATCGTGTGTTGAACAGCTCAGAGAGAAATGAACGGTT 1094
 Qy 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
 Db 1095 TTTCAGCTCTCATTTACTCTCTCAACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1145
 Qy 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
 Db 1146 GAAGGAGACCCGGAAGCTCAAGAGAGATATCCAAAATATCC----- 1187
 Qy 292 ProGlyGluAlaSerCysSerGluThrProLysArgProLysValLysArgIlePro 311
 Db 1188 -----AAGTATATGCAAGAACGACAGAAAGG----- 1214
 Qy 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrGlnAsnSerGlyValArg 331
 Db 1215 -----GAGTCACAGACACTGTTGACAGAAATGATGATGATGATGATGATGATGATG 1259
 Qy 332 ValGlu-----ArgGluLeuAlaAlaGlu 339
 Db 1260 GAGGAATGGGAGAGCCAGAGGACAGTCATCTAGGGCTCATGCTCTACACCTGAGTCA 1319
 Qy 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
 Db 1320 CGAGCTGCTGTCAGGAACCTTCCAGCAGTATCTCTCGCTGGTGAAGACCCAGAGAAAGG 1379
 Qy 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376

Db 1380 GGAGTAAACTTGATGGGAGATTCATTTCTACAGTGTCTGTTGGTAAAGCCTCA 1439
 Qy 377 SerTyrPhe-----AspTTPAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
 Db 1440 GCACACCGAGTGGAGACTGGACACACCATAGCCTGTTCGTAGCCATATTAATTTGGT 1499
 Qy 395 LeuCysPheThrLeuValLeuAlaValPheIleValArgAlaLeuProAlaLeu-GlnPh 414
 Db 1500 TTGTGCTTACATTAATTTACTCTTGGCATTTTCAAGAAAGCATTTGCCAGCTCTTCCAAATC 1559
 Qy 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLe 434
 Db 1560 TCCA--TCACCTTTGGGCTGTGTTTCTTACTTTGCCACAGATATCTTTGTACAGCCTTTT 1616
 Qy 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh 451
 Db 1617 ATGGACCAATAGCATTCATCAATTTTATATCTAGCATATTTGCGGTAGATCCCATG 1676
 Qy 451 eLeuCysIleIle 455
 Db 1677 GATGTTTCTTCTT 1689

RESULT 12

AAT59536

ID AAT59536 standard; cDNA; 1750 BP.

XX AC

XX AC

XX AC

XX 07-MAY-1997 (first entry)

XX Human early onset Alzheimer's disease (EOAD) splice variant gene.

DE Human early onset Alzheimer's disease (EOAD) splice variant gene.

XX Early onset Alzheimer's disease; EOAD; neurodegenerative disease;

XX diagnosis; gene therapy; antisense; ds.

XX Homo sapiens.

XX Key

FH Location/Qualifiers

FT 174..1565

FT /*tag= a

FT variation

FT 248..249

FT /*tag= b

FT /label= VRSQ

FT /note= "splice site"

XX MO9703086-A1.

XX 30-JAN-1997.

XX 26-JUN-1996; 96WO-US11064.

XX 18-JUL-1995; 95US-0001501.

XX 13-JUL-1995; 95US-0001142.

XX (UYSF-) UNIV SOUTH FLORIDA.

XX Hardy JA;

XX WPI; 1997-118980/11.

XX P-PSDB; AAW11840.

XX Early onset Alzheimer's disease gene - useful for diagnosing a

XX pre-disposition to Alzheimer's disease

XX Claim 1; Fig 2; 44pp; English.

XX A full-length cDNA (AAT59536) of the early onset Alzheimer's disease

XX (EOAD) splice variant gene sequence codes for a 463-amino acid

XX polypeptide (AAW11839). Another full-length cDNA (AAT59535) of the

XX EOAD gene codes for a 467-amino acid polypeptide (AAW11840). The 2

XX sequences can be used to generate primers and probes for the

XX diagnosis of predisposition to Alzheimer's disease, esp. EOAD.

CC They can also be used for prodn. of EOAD polypeptides in
 CC transformed host cells, and antisense sequences can be used for
 CC the treatment of EOAD.

XX Sequence 1750 BP; 442 A; 389 C; 430 G; 478 T; 11 other;
 SQ

Alignment Scores:

Score: 991.00 Length: 1750
 Percent Similarity: 65.98% Matches: 215
 Best Local Similarity: 49.43% Mismatches: 103
 Query Match: 41.67% Indels: 45
 DB: 18 Gaps: 8

US-09-043-944-1 (1-461) x AAT59536 (1-1750)

Qy 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIle 51
 Db 351 GAGCAAGATGAGGAAGAAGATGAGAGCTGACATTTGAATATGGCCCAACATGTGATC 410
 Qy 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrIle 71
 Db 411 ATGCTCTTTGTCCTGTGACTCTCTGTCATGTGTGGTGGTGGTACCATTAAGTCAGTC 470
 Qy 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
 Db 471 AGCTTTTATACCGGAAGATGGGCAG---CTAATCTATATACCCCATTCACAGAAATACC 527
 Qy 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
 Db 528 GAGACTGTGGCCAGAGAGCCCTGCACCTCAATCTCGAATGTGCCATCATGATCAGTGC 587
 Qy 112 ValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuIle 131
 Db 588 ATTGTGTGTCAGTATCTCTCTGCTGTCTCTATAAATACAGGTGTCTATAAGTCATC 647
 Qy 132 HisGlyTyrLeuIleValSerSerPheLeuLeuPheLeuPheThrThrIleTyrVal 151
 Db 648 CATGCTGCTTATATATATCATCTCTATGTGTGTGTTCTTTTTCATTCATTACTTG 707
 Qy 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
 Db 708 GGGGAAGTGTTTAAACCTATAACGTTGCTGGGACTACATTACTGTGCACTCCTGATC 767
 Qy 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTyrLysGlyProLeuArgGluGln 191
 Db 768 TGGAAATTTGGTGTGGTGGATGATTTCCATTCTCGAAGGTCCTTCGATCCAG 827
 Qy 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
 Db 828 CAGGCATATCTCATTTATGATGATGCTCATGGCCCTGGTGTGTTATCAAGTACCTCCT 887
 Qy 212 GluTyrThrValTyrPheValLeuPheValIleSerValTyrPheValValAlaValLeu 231
 Db 888 GAATGGACTGCGTGTGCTCATCTGCTGTGATTTTTCAGTATATGATTTAGTGGCTGTTTG 947
 Qy 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
 Db 948 TGTCGGAAGGTCCACTTCGTATGCTGTGTGTAACAGCTCAGGAGAGAAATGAACGCTT 1007
 Qy 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuThrAlaVal 271
 Db 1008 TTTCCAGCTCTCATTTACTCTCTCAACATGGTGTGG-----TTGGTGAATATGCA 1058
 Qy 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
 Db 1059 GAAGGAGACCCGAGAGCTCAAGAGAGATATCCAAAAATTC----- 1100
 Qy 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
 Db 1101 -----AAGTATAATGACAGAAAGCAGACAGAG----- 1127
 Qy 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331

Db 1071 GAAGGAGACCCGGAAGCTCAAGAGAGATCCAAAAATTC----- 1112
 Qy 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
 Db 1113 -----AAGTATAATGCAGAAAGCAGAAAGG----- 1139
 Qy 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331
 Db 1140 -----GAGTCACAGACACTGTTCAGAGAGATCATGATGGCGGTTCAGT 1184
 Qy 332 ValGlu-----ArgGluLeuAlaAlaGlu 339
 Db 1185 GAGGAATGGGAAGCCAGAGGAGCAGTCATCTAGGGCTCATCGCTCTACACCTGAGTCA 1244
 Qy 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
 Db 1245 CGAGCTCTGTCAGGAACCTTCCAGCAGATCTCTGCTGGTGAAGACCAGAGGAAGG 1304
 Qy 357 GlyValLysLeuGlyLeuGlyAspPheIlePheThrSerValLeuLeuGlyLysAlaSer 376
 Db 1305 GGAGTAAACCTGGATTGGGAGATTTCATTTCTACAGTGTCTGGTTGGTAAAGCCTCA 1364
 Qy 377 SerTyrPhe-----AspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuGly 394
 Db 1365 GCAACACCGAGTGGAGCTGGAAACACACCATAGCTGTTCGTAGCCATATTAATGGT 1424
 Qy 395 LeuCysPheThrLeuValLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPh 414
 Db 1425 TTGGCGCTTACATTAATTAATCTCTGCTGCTTCCATTTCAAGAAAGCATTGCCAGCTCTTCCAATC 1484
 Qy 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLe 434
 Db 1485 TCCA---TCACCTTTGGGCTGTGTTCTTCTACTTTGCCACAGATTATCTGTACAGCCTTTT 1541
 Qy 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe 448
 Db 1542 ATGGACCAATTAGCATTCATCAATTTTATATCTAGCATATTT 1584
 RESULT 14
 ABK83912
 ID ABK83912 standard; cDNA; 3056 BP.
 AC ABK83912;
 XX 14-AUG-2002 (first entry)
 DE Human cDNA differentially expressed in granulocytic cells #483.
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 OS Homo sapiens.
 XX
 XX WO200228999-A2.
 FN
 XX
 XX 11-APR-2002.
 PD
 XX
 XX 03-OCT-2001; 2001WO-US30821.
 PF
 XX
 XX 03-OCT-2000; 2000US-237189P.
 PR
 XX
 XX (GENE-) GENE LOGIC INC.
 PA
 XX
 XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 PI
 XX
 XX WPI; 2002-435328/46.
 DR
 XX

PT Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity
 XX
 PS Claim 1; SEQ ID No 483; 114pp; English.
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC or allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating GA; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, ARDS, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, asthma, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease; also bacterial infection, viral infection,
 CC parasitic infection, protozoal infection, fungal infection and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3056 BP; 762 A; 688 C; 740 G; 866 T; 0 other;

Alignment Scores:
 Pred. No.: 2,29e-89 Length: 3056
 Score: 989.50 Matches: 216
 Percent Similarity: 65.62% Conservatives: 76
 Best Local Similarity: 48.54% Mismatches: 105
 Query Match: 41.61% Indels: 48
 DB: 24 Gaps: 9
 US-09-043-944-1 (1-461) x ABK83912 (1-3056)
 QY 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIle 51
 Db 731 GAGCAAGATGAGAGAGAGAGATGAGAGCTGACATTTGAATATATGGCCCAAGCTGTGATC 790
 QY 52 HisLeuPheValProValSerLeuLysMetAlaLeuValValPheThrMetAsnThrIle 71
 Db 791 ATGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
 QY 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
 Db 851 AGCTTTTATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 907
 QY 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
 Db 908 GAGACTGTGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 967
 QY 112 ValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuIle 131

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2003, 07:41:57 ; Search time 2164 Seconds
(without alignments)
708.031 Million cell updates/sec

Title: US-09-043-944-1

Perfect score: 2378

Sequence: 1 MPSTRQQGGGADAETHV.....YYINSLFLFLCTINFSIIS 461

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09043944/runat_11122003_062402_3791/app_query.fasta_1.647
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -LISTS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LISTS=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORMEXT -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09043944 @CGN 1 1 107 @runat_11122003_062402_3791
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	994.5	41.8	2765	9	US-09-785-474-1	Sequence 1, Appli
2	994.5	41.8	2765	9	US-09-785-474-27	Sequence 27, Appli
3	993.5	41.8	2763	13	US-10-293-000-7	Sequence 7, Appli
4	993.5	41.8	2764	13	US-10-116-275-290	Sequence 230, App
5	993.5	41.8	2765	15	US-10-221-454-5	Sequence 5, Appli
6	989.5	41.6	2765	9	US-09-785-474-31	Sequence 31, Appli
7	984.5	41.4	2765	9	US-09-785-474-29	Sequence 29, Appli
8	979.5	41.2	2765	9	US-09-785-474-3	Sequence 3, Appli
9	979	41.2	1392	9	US-09-895-035-13	Sequence 13, Appli
10	979	41.2	1404	9	US-09-754-949-3	Sequence 3, Appli
11	979	41.2	1404	13	US-10-417-422-3	Sequence 25, Appli
12	967	40.7	2285	9	US-09-878-454A-25	Sequence 1, Appli
13	961	40.4	2236	13	US-10-180-781-1	Sequence 8, Appli
14	959	40.3	2236	13	US-10-293-000-8	Sequence 7, Appli
15	959	40.3	2236	15	US-10-221-254-7	Sequence 5, Appli
16	942.5	39.6	1346	9	US-09-754-949-5	Sequence 5, Appli
17	942.5	39.6	1346	13	US-10-417-422-5	Sequence 1045, Ap
18	454	19.1	1362	10	US-09-938-842A-1045	Sequence 14, Appli
19	337.5	14.2	473	13	US-10-180-781-14	Sequence 8, Appli
20	337.5	14.1	624	9	US-09-895-035-8	Sequence 3, Appli
21	272	11.4	493	9	US-09-895-035-3	Sequence 6, Appli
22	259.5	10.9	469	9	US-09-895-035-6	Sequence 2, Appli
23	259.5	10.9	816	9	US-09-895-035-2	Sequence 28131, A
24	259	10.9	562	11	US-09-918-995-28131	Sequence 4373, Ap
25	248	10.4	411	13	US-10-029-386-4373	Sequence 18073, A
26	246	10.3	245	13	US-10-029-386-18073	Sequence 38, Appli
27	236	9.9	2387	13	US-10-180-781-38	Sequence 5, Appli
C	202.5	8.5	356	9	US-09-895-035-5	Sequence 4, Appli
28	195.5	8.2	269	9	US-09-895-035-4	Sequence 454, App
30	186.5	7.8	332	9	US-09-925-299-454	Sequence 412, Ap
31	186.5	7.8	332	11	US-09-818-875-4212	Sequence 4213, Ap
32	178	7.5	121	11	US-09-818-875-4212	Sequence 4216, Ap
C	178	7.5	121	11	US-09-818-875-4213	Sequence 4217, Ap
34	178	7.5	121	11	US-09-818-875-4216	Sequence 4212, Ap
C	178	7.5	121	11	US-09-818-875-4217	Sequence 4213, Ap
36	178	7.5	121	13	US-10-209-787-4212	Sequence 4216, Ap
C	178	7.5	121	13	US-10-209-787-4213	Sequence 4217, Ap
C	178	7.5	121	13	US-10-209-787-4216	Sequence 4220, Ap
C	178	7.5	121	11	US-09-818-875-4220	Sequence 4220, Ap
40	176	7.4	121	11	US-09-818-875-4221	Sequence 4221, Ap
C	176	7.4	121	13	US-10-209-787-4220	Sequence 4221, Ap
42	176	7.4	121	13	US-10-209-787-4221	Sequence 4224, Ap
C	176	7.4	121	11	US-09-818-875-4224	Sequence 4225, Ap
C	165	6.9	121	11	US-09-818-875-4225	
C	165	6.9	121	11	US-09-818-875-4225	

ALIGNMENTS

RESULT 1

US-09-785-474-1
; Sequence 1, Application US/09785474
; Patent No. US20010012626A1

; GENERAL INFORMATION:
; APPLICANT: TANZI, RUDOLPH

; WASCO, WILMA
; TITLE OF INVENTION: Genetic Alterations Related To Familial

; Alzheimer's Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON

; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA: US/09/785, 474

```

; FILING DATE: 20-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/706,344
; FILING DATE: 30-AUG-1996
; APPLICATION NUMBER: 60/003,054
; FILING DATE: 31-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0609.4180002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2765 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 249..1649
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-785-474-1

Alignment Scores:
Pred. No.: 3,17e-98 Length: 2765
Score: 994.50 Matches: 215
Percent Similarity: 65.17% Conservative: 75
Best Local Similarity: 48.31% Mismatches: 107
Query Match: 41.82% Indels: 48
DB: 9 Gaps: 8

US-09-043-944-1 (1-461) x US-09-785-474-1 (1-2765)
QY 32 GInGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIle 51
Db 438 GAGCAGAGATGAGAGAGAGATGAGAGCTGCATTTGAAATATGCGCCACGACATGTGATC 497
QY 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrIle 71
Db 498 ATGCTCTTTGTCCCTGTGACTCTCTGCGATGCTGGTGGTGGTACCATTAAGTCAGTC 557
QY 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
Db 558 AGCTTTTATACCCGGAAGATGGGCAG---CTAATCTATACCCCATTCACAGAAGATACC 614
QY 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
Db 615 GAGACTGTGGCCAGAGACCCCTGCACCTCAATCTGAATGCTGCCCATCATGACGTGTC 674
QY 112 ValValLeuMetThrValLeuLeuLeuValPheTyrLysTyrLysPheTyrLysLeuIle 131
Db 675 ATTGTTGTCACTGACTATCCTCTGCTGGTCTCTGTATAATACAGGTGCTATAGGTCATC 734
QY 132 HisGlyTyrLeuIleValSerSerPheLeuLeuLeuPheLeuPheThrIleTyrVal 151
Db 735 CATGCTGGCTTATTATATCATCTCTAFTGTGTGTCCTCTTTTTCATTCAATTACTTGG 794
QY 152 GInGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
Db 795 GGGGAAGTGTTTAAACCTATAACGTTGCTGTGGACTACATTAATGTCACCTCGTCATC 854
QY 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTrpLysGlyProLeuArgLeuGln 191
Db 855 TGGAAATTTTGGTGTGGTGGGAATGATTTCATTCCTACCTGGAAAGGTCCACTTCGACTCC 914
QY 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
Db 915 CAGGCATATCTCATTATGATTAGTGGCCCTCATGCCCCTGGTGTATCAAGTACCTCCCT 974

```



```
US-09-043-944-1 (1-461) x US-10-116-275-290 (1-2764)
QY 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIle 51
Db 438 GAGCAAGATGAGGAGAGATGAGAGCTGACATTTGAATATATGCGCCAGCAATGTGATC 497
QY 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrIle 71
Db 498 ATGCTCTTTGTCCTGTGACTCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
QY 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
Db 558 AGCTTTTATACCGGAGAGATGGGAG--CTAATCTATACCCCAATTCACAGAGATACC 614
QY 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
Db 615 GAGACTGTGGCCAGAGAGCCCTGCACTCAATCTGAATGCTGCCATCATGATGATGTC 674
QY 112 ValValLeuMetThrValLeuLeuIleValPheTyrIleLysTyrLysPheTyrLysLeuIle 131
Db 675 ATTGTTGTCATGACTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734
QY 132 HisGlyTrpLeuIleValSerPheLeuLeuLeuPheLeuPheThrThrIleTyrVal 151
Db 735 CATGCTGGCTTATATATATATATATATATATATATATATATATATATATATATATAT 794
QY 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
Db 795 GGGGAGTGTGTTAAAGCTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854
QY 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTrpLysGlyProLeuArgLeuGln 191
Db 855 TGGAAATTTGCTGGTGGGGAATGATTTCCATTCACCTGGAAGAGTCCACTTCGACTCCAG 914
QY 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
Db 915 CAGGCATATCTATATATATATATATATATATATATATATATATATATATATATATAT 974
QY 212 GluTrpThrValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeu 231
Db 975 GAATGGAGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034
QY 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
Db 1035 TGTCGGAAGTCCACTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1094
QY 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
Db 1095 TTTCAGCTCTCATTTACTCTCTCAACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145
QY 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
Db 1146 GAAGGAGACCGGAGGCTCAAGAGAGATTCGAAATATCC----- 1187
QY 292 ProGlyGluAlaSerCysSerGluThrProLysArgProLysValLysArgIlePro 311
Db 1188 -----AAGTATAATGCAGAAAGCAGACAGAAAGG----- 1214
QY 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331
Db 1215 -----GAGTCACAGACACTGTTGCAGAGAAATGATGCTGCTGCTGCTGCTGCTGCTGCT 1259
QY 332 ValGlu----- 339
Db 1260 GAGGATGGGAGCCAGAGGAGCAGTCATCTAGGGCTCATCGCTTACACCTGAGTCA 1319
QY 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
Db 1320 CGAGCTGCTGCTCAGGAACCTTCCAGCAGTATCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1379
QY 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
Db 1380 GGAGTAAAACTTGGATGGGAGATTTCAATTTTCTACAGTGTCTGTTGGTAAAGCCCTCA 1439
QY 377 SerTyrPhe-----AspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
Db 1440 GCAACAGCCAGTGGGAGACTGGAAACAACCATAGCTGTTTCGTAGCCATATTAATGCT 1499
QY 395 LeuCysPheThrLeuLeuValLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPhe 414
Db 1500 TTGTGCTTACATTAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1559
QY 414 eProPheSerProAspSerPheThrPheValProAlaGlySerSerProHisLeuLe 434
Db 1560 TCCA---TACCTTTGGGCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1616
QY 434 uHisLysSerLeuLysSerValTyrIleAsnSerLeuPhe-----LeuProPhe 451
Db 1617 ATGGACCAATAGCATTCATCAATTTTATATATCTAGCATATTTGGGGTTAGAAATCCCATG 1676
QY 451 eLeuCysIleLeu 455
Db 1677 GATGTTTCTTCTT 1689

RESULT 5
US-10-221-254-5
; Sequence 5, Application US/10221254
; Publication NO. US20030113811A1
; GENERAL INFORMATION:
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: ASSAY
; FILE REFERENCE: PG3950USW
; CURRENT APPLICATION NUMBER: US/10/221,254
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0005894.1
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (249)..(1652)
US-10-221-254-5
Alignment Scores:
Pred. No.: 4,09e-98 Length: 2765
Score: 993.50 Matches: 216
Percent Similarity: 65.84% Conservative: 77
Best Local Similarity: 48.54% Mismatches: 104
Query Match: 41.76% Indels: 48
DB: 15 Gaps: 9

US-09-043-944-1 (1-461) x US-10-221-254-5 (1-2765)
QY 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIle 51
Db 438 GAGCAAGATGAGGAGAGATGAGAGCTGACATTTGAATATATGCGCCAGCAATGTGATC 497
QY 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrIle 71
Db 498 ATGCTCTTTGTCCTGTGACTCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
QY 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
Db 558 AGCTTTTATACCGGAGAGATGGGAG--CTAATCTATACCCCAATTCACAGAGATACC 614
QY 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
Db 615 GAGACTGTGGCCAGAGAGCCCTGCACTCAATCTGAATGCTGCCATCATGATGATGTC 674
QY 112 ValValLeuMetThrValLeuLeuIleValPheTyrIleLysTyrLysPheTyrLysLeuIle 131
Db 675 ATTGTTGTCATGACTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734
QY 132 HisGlyTrpLeuIleValSerPheLeuLeuLeuPheLeuPheThrThrIleTyrVal 151
Db 735 CATGCTGGCTTATATATATATATATATATATATATATATATATATATATATATATAT 794
QY 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
Db 795 GGGGAGTGTGTTAAAGCTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854
QY 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTrpLysGlyProLeuArgLeuGln 191
Db 855 TGGAAATTTGCTGGTGGGGAATGATTTCCATTCACCTGGAAGAGTCCACTTCGACTCCAG 914
QY 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
Db 915 CAGGCATATCTATATATATATATATATATATATATATATATATATATATATATATAT 974
QY 212 GluTrpThrValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeu 231
Db 975 GAATGGAGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034
QY 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
Db 1035 TGTCGGAAGTCCACTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1094
QY 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
Db 1095 TTTCAGCTCTCATTTACTCTCTCAACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145
QY 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
Db 1146 GAAGGAGACCGGAGGCTCAAGAGAGATTCGAAATATCC----- 1187
QY 292 ProGlyGluAlaSerCysSerGluThrProLysArgProLysValLysArgIlePro 311
Db 1188 -----AAGTATAATGCAGAAAGCAGACAGAAAGG----- 1214
QY 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331
Db 1215 -----GAGTCACAGACACTGTTGCAGAGAAATGATGCTGCTGCTGCTGCTGCTGCTGCT 1259
QY 332 ValGlu----- 339
Db 1260 GAGGATGGGAGCCAGAGGAGCAGTCATCTAGGGCTCATCGCTTACACCTGAGTCA 1319
QY 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
Db 1320 CGAGCTGCTGCTCAGGAACCTTCCAGCAGTATCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1379
QY 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
Db 1380 GGAGTAAAACTTGGATGGGAGATTTCAATTTTCTACAGTGTCTGTTGGTAAAGCCCTCA 1439
QY 377 SerTyrPhe-----AspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
Db 1440 GCAACAGCCAGTGGGAGACTGGAAACAACCATAGCTGTTTCGTAGCCATATTAATGCT 1499
QY 395 LeuCysPheThrLeuLeuValLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPhe 414
Db 1500 TTGTGCTTACATTAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1559
QY 414 eProPheSerProAspSerPheThrPheValProAlaGlySerSerProHisLeuLe 434
Db 1560 TCCA---TACCTTTGGGCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1616
QY 434 uHisLysSerLeuLysSerValTyrIleAsnSerLeuPhe-----LeuProPhe 451
Db 1617 ATGGACCAATAGCATTCATCAATTTTATATATCTAGCATATTTGGGGTTAGAAATCCCATG 1676
QY 451 eLeuCysIleLeu 455
Db 1677 GATGTTTCTTCTT 1689
```

Db 675 ATTGTTGTCAGTATCCCTCGTGGTCTGTGTAATAATACAGGTGCTATAAGTCAATC 734
Qy 132 HisGlyTrpLeuIleValSerSerPheLeuLeuLeuPheLeuPheThrThrIleVal 151
Db CATCCCTGGGCTTATATATCATCTCTATGTTGCTGTTCTTTTTCATTCATTTACTTG 794
Qy 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
Db GGGGAAGTGTAAACCTATAACGTTGCTGTGGACTACATTAATGCTGCTCTCTGATC 854
Qy 172 GlyAsnTyrGlyValLeuGlyMetCysIleHisTrpLysGlyProLeuArgLeuGln 191
Db TGGATTTTGGTGGTGGGATGATTCCTCATCTGGAAGGTCACITTCGACTCAG 914
Qy 192 GlnPheTyrIleLeuThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
Db CAGGCATATCTCATTAATAGTAGCCCTCATGCCCTGGTGTATCAAGTACCTCCCT 974
Qy 212 GluTrpThrValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeu 231
Db GAATGGACTGGTGGCTCATCTTGGCTGTGATTCAGTATATGATTTAGTGGCTGTTTG 1034
Qy 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
Db TGTCGGAAGGTCACATTCGTATGCTGTTGAAACAGCTCAGAGAGAAATGAAACGCTT 1094
Qy 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
Db TTTCCAGCTCTCATTTACTCTCAACATATGTTGG-----TTGGTGAATATGGCA 1145
Qy 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
Db GAAGGAGACCCGGAGCTCAAGAGAGATATCCAAAATATCC----- 1187
Qy 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
Db AAGTATAATGCAGAAAGCACAGAAAG----- 1214
Qy 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331
Db GAGTCAAGACACTGTTGCAGAGATGATGCGGGGTTTCAGT 1259
Qy 332 ValGlu-----ArgGluLeuAlaGlu 339
Db GAGAAATGGGAAGCCAGAGGACAGTCATCTAGGGCTCATCTCTACACCTGAGTCA 1319
Qy 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
Db CGAGCTGCTGCCAGAACTTTCAGCAGTATCTCGTGTGAAGACCCAGAGAAAGG 1379
Qy 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
Db GGAGTAAACTTGATTTGGGAGATTTTCATTTCTACAGTGTCTGTTGTTAAAGCCTCA 1439
Qy 377 SerTyrPhe-----AspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
Db GCAACACCGAGTGGAGACTGGACACACACCATAGCTGTTCTGTAGCATATTAATTTGT 1499
Qy 395 LeuCysPheThrLeuValLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPhe 414
Db TTGTGCTTACATTAATTAATCTTCCATTTTCAAGAAAGCATTGCCAGCTCTTCCCAATC 1559
Qy 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLe 434
Db TCCA---TCACCTTTGGGCTGTTTCTTACTTGTCCACAGATATATCTTTGACAGCCTTTT 1616
Qy 434 uHisLysSerLeuLysSerValTyrIleAsnSerLeuPhe-----LeuProPhe 451
Db ATGGACCAATTAGCATTCATATTTTATATCTAGCATATTTTCGGGTGAGATATCCCATG 1676
Qy 451 eLeuCysIleIle 455
Db GATGTTTCTTCTT 1689

RESULT 6

US-09-785-474-31
; Sequence 31, Application US/09785474
; Patent No. US20010012626A1
; GENERAL INFORMATION:
; APPLICANT: TANZI, RUDOLPH
; TITLE OF INVENTION: Genetic Alterations Related To Familial
; Alzheimer's Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/785,474
; FILING DATE: 20-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/706,344
; FILING DATE: 30-AUG-1996
; APPLICATION NUMBER: 60/003,054
; FILING DATE: 31-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0609.4180002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2765 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 249..1649
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-785-474-31

Alignment Scores:
Pred. No.: 1,12e-97 Length: 2765
Score: 989.50 Matches: 214
Percent Similarity: 64.94% Conservative: 75
Best Local Similarity: 48.09% Mismatches: 108
Query Match: 41.61% Indels: 48
DB: 9 Gaps: 8

US-09-043-944-1 (1-461) x US-09-785-474-31 (1-2765)

Qy 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIle 51
Db GAGCAAGATGAGGAGAGAGATGAGAGCTGACATGGAATATGCGCCAGCATGTGATC 497
Qy 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrIle 71
Db ATGCTCTTGTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 557
Qy 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
Db AGCTTTTATACCCGGAAGGATGGGAG---CTAATCTATACCCCATTCACAGAAGATACC 614

QY	92	AspSerIleValGluIysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal	111
DB	615	GAGACTGTGGCCAGAGAGCCCTGCATCAATTCTGAATGCTGCCATCATGATCATGCTGC	674
QY	112	ValValLeuMetThrValLeuLeuLeuValPheTyrIysTyrIysPheTyrIysLeuIle	131
DB	675	ATTGTTGCTCATGACTATCCCTCCCTGGTGGTCTGTATATAATACAGAGGCTATTAAGGTCATC	734
QY	132	HisGlyTyrLeuIleValSerSerPheLeuLeuPheLeuPheThrThrIleTyrVal	151
DB	735	CATGCCCTGGCTATTATATATCATCTCTATTGCTGCTGCTCTTTTTCATTTCAATTTACTTGG	794
QY	152	GlnGluValLeuIysSerPheAspValSerProSerAlaLeuValLeuPheGlyLeu	171
DB	795	GGGGAAGTGTATAAACCCTAATGCTGTGTGGACTACATTAATCTGTGTGCATCCCTCGATC	854
QY	172	GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTyrIysGlyProLeuArgLeuGln	191
DB	855	TGGAAATTTGGTGTGGTGGAAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAG	914
QY	192	GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleIysTyrLeuPro	211
DB	915	CAGGCATATCTCATTTATGATAGTGGCCCTCATGGCCCTGGTGTATTAACAAGTACCTCCCT	974
QY	212	GluTyrThrValTyrPheValLeuPheValIleSerValTyrAspLeuValAlaValLeu	231
DB	975	GAATGGACTCGGTGGCTCATCTTGGCTGTGATTTCAATATGATTTAGTGGCTGTTTG	1034
QY	232	ThrProIysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle	251
DB	1035	TGTCGGAAGGTCCACTTCATATGCTGGTTGAAACAGCTCAGGAGAAATGAACGCCTT	1094
QY	252	PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal	271
DB	1095	TTTCCAGCTCTCATTTACTCTCTCAACATGGTGGT-----TTGGTGAATATGGCA	1145
QY	272	GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe	291
DB	1146	GAAGGAGACCCGGAAGCTCAAGAGAGAGTATCCAAAATTC-----	1187
QY	292	ProGlyGluAlaSerCysSerSerGluThrProIysArgProIysValIysArgIlePro	311
DB	1188	-----AAGCATATGCA	1199
QY	312	GlnIysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg	331
DB	1200	GAAGAGCACAGAAAGGGAGTCAACAGACACTGTTTCAGAGAATGATGATCGCGGGTTCAGT	1259
QY	332	ValGlu-----ArgGluLeuAlaAlaGlu	339
DB	1260	GAGAAATGGGAAGCCAGAGGACAGTCACTAGGCGCTCATCGCTACACCTGAGTCA	1319
QY	340	ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluGluArg	356
DB	1320	CGAGCTGTGTCGAGGAACCTTCCAGCAGTATCTCGCTGGTGAAGACCCAGAGAAAGG	1379
QY	357	GlyValIysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyIysAlaSer	376
DB	1380	GGAGTAAACCTTGGATTGGAGAGATTTCATTTTCACAGTGTTCUGGTGTGAAGCCCTCA	1439
QY	377	SerTyrPhe-----AspTyrPAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly	394
DB	1440	GCAACAGCCAGCTGGAGACTGGAAACACACACACATAGCTGTTTCGTAGCCATATTAAATGGT	1499
QY	395	LeuCysPheThrLeuValLeuValAlaValPheIysArgAlaLeuProAlaLeu-GlnPhe	414
DB	1500	TTGTGCCCTTACATATTACTCTCTGCCATTTCAAGAAAGCATTTGCAGACTCTTCCAATC	1559
QY	414	eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLe	434
DB	1560	TCGCA---TCACCTTTGGGCTGTGTTCTTCTACTTTGCTGCACAGATTTATCTGTACAGCCTTTT	1616

```

Qy 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh 451
Db 1617 ATGGACCAATTCAGATCCCATCAATTTTATATAGCATATTTTCGGGTTAGAAATCCCATG 1676
Qy 451 eLeuCysIleIle 455
Db 1677 GATGTTCTCTT 1689

RESULT 7
US-09-785-474-29
; Sequence 29, Application US/09785474
; Patent No. US20010012626A1
; GENERAL INFORMATION:
; APPLICANT: TANZI, RUDOLPH
; TITLE OF INVENTION: Genetic Alterations Related To Familial
; Alzheimer's Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09785,474
; FILING DATE: 20-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/706,344
; FILING DATE: 30-AUG-1996
; APPLICATION NUMBER: 60/003,054
; FILING DATE: 31-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0609.4180002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2765 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 249..1649
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-785-474-29

```

QY 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrIle 71
DB 498 ATGCTCTTTGTCCTGCTGACTCTCGCAGTGGGGTGGTGGCTGACCTAAAGTCAGTC 557
QY 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
DB 558 AGCTTTTATACCGGAGGATGGGAG--CTAATCTATACCCCATTCACAGAAATACC 614
QY 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
DB 615 GAGACTGTGGGCGAGAGCCCTGCACCTCAATTCGAATGCTGCCCATCATGATGATGTC 674
QY 112 ValValLeuMetThrValLeuLeuLeuValPheTyrLysTyrLysPheTyrLysLeuIle 131
DB 675 ATTGTGTCATGACTATCTCTGCTGGTGTCTGTATAATACAGGTGCTATAAGGTATC 734
QY 132 HisGlyTyrLeuIleValSerSerPheLeuLeuPheLeuPheThrThrIleTyrVal 151
DB 735 CATGCTGGCTTATTATATCATCTCTATTGTGCTGTTCTTTTTCATTTCATTACTTG 794
QY 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
DB 795 GGGGAAGTGTAAACACCTATAACGTTGCTGGACTACATTACTGTTGCACCTCCTGATC 854
QY 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTyrLysGlyProLeuArgGluIn 191
DB 855 TGGAAATTTGTGTGGTGGATGATGATTCCTACTGGAAGGTCACCTTCGACTCCAG 914
QY 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
DB 915 CAGGCATATCTCATTTAGTATGATGCTCCCTCATGGCCCTGGTGTATTCAAGTACCTCCT 974
QY 212 GluTyrThrValTyrPheValLeuPheValIleSerValTyrAspLeuValAlaValLeu 231
DB 975 GAATGGACTGGTGGCTCATCTGCTGTGATTTTCAGTATATGATTTAGTGGCTGTTTG 1034
QY 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
DB 1035 TGCTGGAAGTCCACTTCGTATGCTGTGTGAAACGCTCAGGAGAAATGAACGCTT 1094
QY 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
DB 1095 TTTCAGCTCTCATTTACTCTCAACAATGGTGG-----TTGGTGAATATGGCA 1145
QY 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
DB 1146 GAAGAGACCCGGAGCTCAAGAGGAGATATCCAAATTC----- 1187
QY 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
DB 1188 -----AAGCAATAATGCA 1199
QY 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrGlnAsnSerGlyValArg 331
DB 1200 GAAAGCAGAGAAAGGGAGTCAAGACACTCTTGACAGAAATGATGGGGGTTTCAGT 1259
QY 332 ValGlu-----ArgGluLeuAlaAlaGlu 339
DB 1260 GAGGAATGGGAGCCGAGAGGAGCAGTCACTAGGGCCCTCATCGCTTCACACCTGAGTCA 1319
QY 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
DB 1320 CGAGCTGCTGCCAGAACTTCCAGCAGTATCTCGCTGGTGGAGACCCAGAGGAAGG 1379
QY 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
DB 1380 GGAGTAAAACTTGGATTGGGAGATTTCATTTCTACAGTGTCTGTTGGTTAAAGCCTCA 1439
QY 377 SerTyrPhe-----AspTyrAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
DB 1440 GCAACAGCCAGTGGAGACTGGAAACACACCATAGCCTGTTTCGTAGCCATTAATGTT 1499
QY 395 LeuCysPheThrLeuValLeuLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPh 414

DB 1500 TTGTGCTTACATTATTACTCTTCCATTTCAGAAAGCATTCGCCAGCTCTTCCAATC 1559
QY 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLe 434
DB 1560 TCCA---TCACCTTGGGCTTGTTCCTACTTTCACAGATATATCTTGTACAGGCTTTT 1616
QY 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh 451
DB 1617 ATGGACCAATTAGCATTCATCAATTTTATATATAGCATATTGGGGTTAGAAATCCCATG 1676
QY 451 eLeuCysIleIle 455
DB 1677 GATGTTCTTCTT 1689
RESULT 8
US-09-785-474-3
; Sequence 3, Application US/09785474
; Patent No. US20010012626A1
; GENERAL INFORMATION:
; APPLICANT: TANZI, RUDOLPH
; WASCO, WILMA
; TITLE OF INVENTION: Genetic Alterations Related To Familial
; Alzheimer's Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/785,474
; FILING DATE: 20-Feb-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/706,344
; FILING DATE: 30-AUG-1996
; APPLICATION NUMBER: 60/003,054
; FILING DATE: 31-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0609.4180002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2765 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 249..1649
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-785-474-3
Alignment Scores:
Pred. No.: 1,41e-96 Length: 2765
Score: 979.50 Matches: 213
Percent Similarity: 64.72% Conservative: 75
Best Local Similarity: 47.87% Mismatches: 109
Query Match: 41.19% Indels: 48
DB: 9 Gaps: 8

US-09-878-454A-25

Alignment Scores:
 Pred. No.: 2,42e-95 Length: 2285
 Score: 967.00 Matches: 220
 Percent Similarity: 61.95% Conservative: 60
 Best Local Similarity: 48.67% Mismatches: 94
 Query Match: 40.66% Indels: 78
 DB: 9 Gaps: 12

US-09-043-944-1 (1-461) x US-09-878-454A-25 (1-2285)

QY	2	ProSerThrArgargGlnGlnGluGly-----GlyGlyAlaAspAlaGluThrHisThr	19
DB	441	CCACGCCGCGCTCCTGCCAGAGGGCAGCAGGCGCCAGAGGATGAGAG--AATACT	497
QY	20	ValTyGlyThrAsnLeuIleThrAsnArgAsnSerGlnGluAspGluAsn-----	36
DB	498	GCCACGTGGAGAGCCAGGAGACGAGGAGCGGTGAGGAGGACCTGACCGGTATGTC	557
QY	37	-----ValValGluGluAlaGluLeuLysTyTrGly	46
DB	558	TGTAGTGGGTTCCCGGGCGGCCCGCAGGCGCTGGAGGAAGAGCTGACCTCAATACCGA	617
QY	47	AlaSerHisValIleHisLeuPheValProValSerLeuCysMetAlaLeuValValPhe	66
DB	618	GCGAAGCATGTGATCATGCTGTGTGTCCTGTCACCTGTGTCATGATCGTGGTGTAGCC	677
QY	67	ThrMetAsnThrIleThrPheTyTrSerGlnAsnAsnGlyValArgHisLeuLeuSerHisPro	86
DB	678	ACCATCAAGTCTGTGGCTTCTACACAGAGAAATGGACAG--CTCATCTACACGCCCA	734
QY	87	PheValargGluThrAspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeu	106
DB	735	TTCACTGAGACACACCCCTCGGTGGCGCAGCGCTCTCTCAACTCGGTGTGAACACCCCTC	794
QY	107	ValMetLeuCysValValValLeuMetThrValLeuLeuIleValPheTyTrLys	126
DB	795	ATCATCATGCGTCTCGTGGTATTGACCATCTTCTTGGTGGTGTCTCAAGTATCCGC	854
QY	127	PheTyTrLysLeuIleHisGlyTyTrLeuIleValSerPheLeuLeuLeuPheLeuPhe	146
DB	855	TGCTACAAGTTCACTCATGGCTGGTGTGATCATGTCTTCACTGAAGTGTCTGTCTCTTC	914
QY	147	ThrThrIleTyTrValGlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeu	166
DB	915	ACCTATATCTACCTTGGGGAAGTCTCAAGACCTACAATGTGGTCATGGACTACCCACC	974
QY	167	ValLeuPheGlyLeuGlyAsnTyTrGlyValLeuGlyMetMetCysIleHisTyTrLys	186
DB	975	CTTCTCTGCTGACTGTCTGGAACCTTCGGGGGAGTGGGATGTTGTGTCATCCACTGGAAGGC	1034
QY	187	ProLeuArgLeuGlnGlnPheTyTrLeuIleThrMetSerAlaLeuMetAlaLeuValPhe	206
DB	1035	CCTCTGGTGTGCAGAGCGCTACCTCATCATGATCAGTGGCTCATGGCCCTTAGTGTTC	1094
QY	207	IleLysTyTrLeuProGluTrpThrValTrpPheValLeuPheValIleSerValTrpAsp	226
DB	1095	ATCAAGTACTCTCCAGAGTGTTCGGGTGCGGTGCATCTCTGGCGCCATCTCTGTGTATGAT	1154
QY	227	LeuValAlaValLeuThrProLysGlyProLeuArgTyTrLeuValGluThrAlaGlnGlu	246
DB	1155	CTCTGGTGTGTCTGTGTCTCCAAAGGGCCCTCTGAGAAATGCTGGTAGAACTGCCAGGAG	1214
QY	247	ArgAsnGluProIlePheProAlaLeuIleTyTrSerSerGlyValIleTyTrProTyVal	266
DB	1215	AGAAATAGAGCCATATTCCTGCCCTGATATACTCATCTGCCATGGTGTGGACGGTGGC	1274
QY	267	LeuValThr-----AlaValGluAsnThrThrAspProArg---	278
DB	1275	ATGGCGAAGTGGACCCCTCTCTCAGGGTGGCTCCAGCTCCCTACGACCCGGAGATG	1334
QY	279	GluProThrSerSerAspSerAsnThrSerThrAlaPheProGlyGluAlaSerCysSer	298

QY 378 ---TyrPheAspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuLeuGlyLeuVal 396
 Db 1508 GGCAGCGGGAGCTGGATATACACGCTGGCTTCTTCTGGCCATCCTCATTTGGCTTGTGT 1567
 QY 397 PheThrLeuValLeuLeuAlaValPheLysArgAlaLeuProAlaLeuGln----- 413
 Db 1568 CTGACCCCTCTCTGTCTGTGTGTCTTCAAGAAGCGCTGCCCGCCCTCCCATCTCCATC 1627
 QY 414 -----PheProPheSerProAspSer 420
 Db 1628 ACGTTGGGCTCATCTTTTACTTCTCCAGGACAAC 1663

RESULT 15
 US-10-221-254-7
 ; Sequence 7, Application US/10221254
 ; Publication No. US20030113811A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GLAXO GROUP LIMITED
 ; TITLE OF INVENTION: ASSAY
 ; FILE REFERENCE: PG3950USW
 ; CURRENT APPLICATION NUMBER: US/10/221,254
 ; PRIORITY FILING DATE: 2002-09-10
 ; PRIOR APPLICATION NUMBER: 0005894.1
 ; PRIORITY FILING DATE: 2000-03-10
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 2236
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (368)..(1714)
 ; US-10-221-254-7

Alignment Scores:
 Pred. No.: 1,76e-94 Length: 2236
 Score: 959.00 Matches: 219
 Percent Similarity: 61.73% Conservative: 60
 Best Local Similarity: 48.45% Mismatches: 95
 Query Match: 40.33% Indels: 78
 DB: 15 Gaps: 12

US-09-043-944-1 (1-461) x US-10-221-254-7 (1-2236)

QY 2 ProSerThrArgArgGlnGlnGly-----GlyGlyAlaAspAlaGluThrHisThr 19
 Db 443 CCCACGCCGCTCTCTGCCAGAGGCGCAGGCGGCCCGCAGGAGTGGAGAG---AACACT 499
 QY 20 ValTyrGlyThrAsnLeuIleThrAsnArgAsnSerGlnGluAspGluAsn----- 36
 Db 500 GCCCAGTGGAGAGCCAGGAGAACGAGGAGACGGTGTGAGGAGGACCTGACCGCTATGTC 559
 QY 37 -----ValValGluAlaGluLeuLysTyrGly 46
 Db 560 TGTAGTGGGTTCGCGGCGCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 619
 QY 47 AlaSerHisValIleHisLeuPheValProValSerLeuCysMetAlaLeuValValPhe 66
 Db 620 GCGAAGCAGCGATCATGCTCTTTGCGCTGCTCATCTGTCATGATCGTGGTGGTAGCC 679
 QY 67 ThrMetAsnThrIleThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisPro 86
 Db 680 ACCATCAAGTCTGTGGCTTCTTACACAGAGAAGATGGACAG---CTCATCTACACGACA 736
 QY 87 PheValArgGluThrAspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeu 106
 Db 737 TTCACTAGGACACACCCCTCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 796
 QY 107 ValMetLeuCysValValValLeuMetThrValLeuLeuValPheTyrLysTyrLys 126
 Db 797 ATCATGATCAGCGTCTGTGGTATGACCATCTTCTTGTGGTGTCTTACAGTACCGC 856

QY 127 PheTyrLysLeuIleHisGlyTrpLeuIleValSerSerPheLeuLeuLeuPheLeuPhe 146
 Db 857 TGTACAAAGTTCATCCATGGCTGTGTGTATCATGTCTTCACTGATGCTGCTGTCTTCTTC 916
 QY 147 ThrThrIleTyrValGlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeu 166
 Db 917 ACCTATATCTCTTGGGAGAGTGTCAAGACCTACAATGTGGCCATGAGCATACCCACC 976
 QY 167 ValLeuPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMetCysIleHisTyrLysGly 186
 Db 977 CTCTTGTGACTGTCTGGAACCTTCGGGAGTGGGAGTGGTGTGTCATCCACTGGAGGGC 1036
 QY 187 ProLeuArgLeuGlnGlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPhe 206
 Db 1037 CCTCTGGTCTGCAGCAGCCTACCTCATCATGATCAGTGGCTCATGCCCTAGTGTTC 1096
 QY 207 IleLysTyrLeuProGluTrpThrValTrpPheValLeuPheValIleSerValTrpAsp 226
 Db 1097 ATCAAGTACCTCCAGAGTGTCCCGTGGGTGCATCTCTGGGCGCATCTCTGTGTATGAT 1156
 QY 227 LeuValAlaValLeuThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGlu 246
 Db 1157 CTCGTGGCTGTGTGTCTCCCAAGGGCTCTGAGATCTCTGTAGAACTGCCCAGGAG 1216
 QY 247 ArgAsnGluProIlePheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrVal 266
 Db 1217 AGAAATGAGCCCATATTCCTGCTGATATATCTCATCTGCCATGTGTGGACGGTTGGC 1276
 QY 267 LeuValThr-----AlaValGluAsnThrThrAspProArg--- 278
 Db 1277 ATGGCGAAGCTGGACCCCTCTCTCAGGGTGGCTTCCAGCTCCCTACGCCGCGAGATG 1336
 QY 279 GluProThrSerSerAspSerAsnThrSerThrAlaPheProGlyGluAlaSerCysSer 298
 Db 1337 GAAGAAGACTCTCTATGACAGTCTT-----GGGAG----- 1366
 QY 299 SerGluThrProLysArgProLysValLysArgIleProGlnLysValGlnIleGluSer 318
 Db 1367 -----CCTTCATACCCCGAGTCTTTGAGCCTCC----- 1396
 QY 319 AsnThrThrAlaSerThrThrGlnAsnSerGlyValValArgValGluArgGluAlaAla 338
 Db 1397 -----TTGACTGGCTACCCAGGAGGAGCTGAG----- 1426
 QY 339 GluArgProThrValGlnAspAlaAsnPheHisArgHisGluGluGluArgGlyVal 358
 Db 1427 -----GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1447
 QY 359 LysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSerSer--- 377
 Db 1448 AAGCTTGGCTCGGGAGCTTCACTTCTTACATGTGTGTGGTGGGAGGAGGAGGAGGAGGAG 1507
 QY 378 ---TyrPheAspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGlyLeuCys 396
 Db 1508 GGCAGCGGGAGCTGGAATACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1567
 QY 397 PheThrLeuValLeuLeuAlaValPheLysArgAlaLeuProAlaLeuGln----- 413
 Db 1568 CTGACCCCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1627
 QY 414 -----PheProPheSerProAspSer 420
 Db 1628 ACGTTGGGCTCATCTTTTACTTCTCCAGGACAAC 1663

Search completed: December 11, 2003, 10:18:31
 Job time : 2190 secs

Result No.	Query #			DB	ID	Description
	Score	Match	Length			
1	2052.5	86.3	4137	4	US-08-832-867-1	Sequence 1, Appl
2	994.5	41.8	2765	3	US-08-706-344C-1	Sequence 1, Appl
3	994.5	41.8	2765	3	US-08-706-344C-27	Sequence 27, Appl
4	993.5	41.8	1488	2	US-08-875-972-3	Sequence 3, Appl
5	993.5	41.8	2764	3	US-08-923-454A-9	Sequence 9, Appl
6	993.5	41.8	2764	4	US-08-832-867-2	Sequence 2, Appl
7	993.5	41.8	2765	3	US-08-888-077A-1	Sequence 1, Appl
8	993.5	41.8	2791	2	US-08-967-101-133	Sequence 133, App
9	993.5	41.8	2791	2	US-08-592-541-133	Sequence 133, App
10	993.5	41.8	2791	3	US-09-124-698-133	Sequence 133, App
11	993.5	41.8	2791	3	US-09-121-480-133	Sequence 133, App
12	993.5	41.8	2791	4	US-09-124-523-133	Sequence 133, App

1

	Oy	18	MetProSerThrArgGlnGlnGlnGlycylalaspalaGluThrHis-----18
	Oy	18	ATGCCTTCCACAGGAGACAACAGAGGGCGGAGTGTCAGATGCCGAACACATGTAAGT101
Ddb	Oy	18	-----18

Db 1020 TATTAGACATTTATTTTCTCAAGAACTAAATTTGTTAAATTTGCTACAAATGCTTGT 1079
 QY 19 -----ThrValTyrGlyThrAsnLeuIleThrAsnArgAsnSerGlnGluAspGluAsnVa 37
 Db 1080 TCAGACCGTTTACGGTACAAATCTGATACAAATCGGAATACCGAAGACGAAATGT 1139
 QY 37 ValGluGluAlaGluLeuLysTyrGlyValAsnHisValIleHisLeuPheValProVa 57
 Db 1140 TGTGAAGAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATCTATTTTGGCCGT 1199
 QY 57 IserLeuCysMetAlaLeuValValPheThrMetAsnThrIleThrPheTyrSerGlnAs 77
 Db 1200 GTCACTATGTCATGGCTCTGGTTGTTTATACATGAACACGATTACGTTTATAGTCAAAA 1259
 QY 77 nAsnGlyArgHisLeu -----LeuSer ----- 82
 Db 1260 CAATGGAAGGCATTTGTAAGTTTCTAAAGAAATGATTCATTAATAATTTTGAATTTGTT 1319
 QY 83 ----- 84
 Db 1320 TTATCAATTTGTCATCTGTGCAATCGCACTCTTTGTGTCAGTGCATAATAATTTTGTGTCAGT 1379
 QY 84 ----- 84
 Db 1380 GCAAAATAATTTTGGTCAGTCATGTTGATATATGTCAGTGCATTTGCAAGTCTGAGCT 1439
 QY 84 ----- 84
 Db 1440 TTAACATTTTTCGTGTTTAAATTTTACTCAATTTTCTATCAATTTTCTTTTGGAAAAA 1499
 QY 84 ----- 84
 Db 1500 GTTGAAGATTTTACTCTGGAATTTTCGAAATAAACTGTAATTTGGAATAATCAATCAACACA 1559
 QY 85 -----His-PropheValArgGluThrAspSerIleValG 96
 Db 1560 AACTTTGAATTTTTCAGACTATACACTCTTTTGTCTCGGAAACACAGACAGTATCGTTG 1619
 QY 96 LuLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysValValValLeuMet 116
 Db 1620 AGAAGGATTTGATGTCACCTGGAAATGCTCTGTCATGTTGTCGTGCTGTTCTGATGA 1679
 QY 116 hrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuIleHisGlyTrpLeuI 136
 Db 1680 CAGTTCTGCTGATTTGTTTCTATAAATCAAGTTTATAAGCTTATTCATGGAATGGCTTA 1739
 QY 136 leValIserSerPheLeuLeuPhePheThrThrIleTyrValGln ----- 152
 Db 1740 TTGTCAGCAGTTTCTCTCTCTTTTCTTCTTATTCACACAAATCTATGTCAGTAAGTTGATA 1799
 QY 153 -----GluValLeuLysSerPheAspValS 161
 Db 1800 TATTACTATTCTCAFAAAATATCAATGTTGAGAGAAGTTCTGAAAAGTTTCGATGTGT 1859
 QY 161 exProSerAlaLeuLeuValLeuPheGlyLeuGlyAsnTyrGlyValLeuGlyMetMetC 181
 Db 1860 CTCCAGCGCACTATTGTTGTTTGTGACTCGGTAACATAAGAGTTCTCGGAATGATGT 1919
 QY 181 yIleHisTrpLysGlyProLeuArgLeuGlnGlnPheTyrLeuIleThrMetSerAlaL 201
 Db 1920 GTATACATTTGGAAGGTCATTCGCTGCAACAGTTCTACCTTATTAACAATGCTGTCAC 1979
 QY 201 euMetAlaLeuValPheIleLysTyrLeuProGluTrpThrValTrpPheValLeuPheV 221
 Db 1980 TAAATGCTCTGGTCTTTATCAAGTACCTACCAAGATGAGCTGTGTTGTTGCTGTTG 2039
 QY 221 alIleSerValTrpAspLeuValAlaValLeuThrProLysGlyProLeuArgTyrLeuV 241
 Db 2040 TTATCTCGGTTTGGATCTGTTGCTGCTCACACCAAAAGACCAATTGAGATATTGG 2099
 QY 241 aIGluThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSer ----- 258

Db 2100 TGGAAACTGCACAGGAGAGAAAAAGCCAAATTTTCCCGCGCTGATTTATTCTGTGTAAGT 2159
 QY 259 -----Ser-Gly 260
 Db 2160 TTCCTAATTTATGGAATTAATATTCATGACGTTTCAAATTTCTAAACACATTTTCAGCTGA 2219
 QY 261 ValIleTyrProTyrValLeuValThrAlaValGluAsnThrThrAspProArgGluPro 280
 Db 2220 GTCATCTATCCCTACGTTCTTGTACTGAGTTGAAACACGACAGACCCCGTGAACCG 2279
 QY 281 ThrSerSerAspSerAsn ----- 286
 Db 2280 ACCTGCTCAGACTCAATATAGTCAGTATCACCTAAATTTTTCGAATTTTTCACAAAAC 2339
 QY 287 -----Thr-SerThrAlaPheProGlyGluAlaSerCysSerSerGluThrProLysAr 304
 Db 2340 AATTTTCAGCTTCTACAGCTTTTCTGGAGAGCGGAGTTGTTTCATCTGAAACGCCAAACG 2399
 QY 304 gProLysValLysArgIleProGlnLysValGlnIleGluSerAsnThrThrAlaSerTh 324
 Db 2400 GCCAAAAGTGAACGAATTCCTCAAAAAGTGCAAAATCGAATCGAATCTACAGCTTCAAC 2459
 QY 324 rThrGlnAsnSerGlyValArgValGluArgGluLeuAlaAlaGluArgProThrValG 344
 Db 2460 GACAAAAACTCTGGAGTAAGGTGAGCGAGCTAGCTCTGAGAGACCAACTGTAC 2519
 QY 344 nAspAlaAsnPheHisArgHisGluGluGluArg ----- 356
 Db 2520 AGACGCCAATTTTTCACAGGACGAGAGAGAGAG -TGAGTGAAAAACGTCGTGAAAAA 2578
 QY 356 ----- 356
 Db 2579 GGGCAAAAGGGGATGATTTTTCGAAAATTTTACTCGAACTTCTCCTCACTTCTAACTCAAAT 2638
 QY 356 ----- 356
 Db 2639 GTTTTTTCTTGACAGCACAAAACGAAAATATTGCCGCTCTACGTTCCGTTATCGAAATATCCC 2698
 QY 356 ----- 356
 Db 2699 CCTGCAATTTTCATTCGTTTTTTTTTCACTGTTTCAATTTTCTCAACTTTTGAAGAGCA 2758
 QY 356 ----- 356
 Db 2759 ATGCCGCCCTCAGCTGAAATATATTTGTTCAATTTAAAGTTCAAAACCTTTTCAGTTAAT 2818
 QY 356 ----- 356
 Db 2819 AGATTCAAGAAAGATCTCAATAAACTTGCAAGCTTGCCACTTGCCGTAGTCACGAAAAA 2878
 QY 356 ----- 356
 Db 2879 AAGGATTTCTCAATAAACCTCTGTTTCATATTTTTTTTAAACAATAATTTTTTCACTCTTC 2938
 QY 357 -----GlyValLysLeuGlyLeuG 363
 Db 2939 ATATTTTGATATGTTTGGCAACAAAAAATGATTGAGGAGGTGTGAAACTTGTCTGGG 2998
 QY 363 YAspPheIlePheTyrSerValLeuLeuGlyLysAlaSerSerTyrPheAspTrpAsnTh 383
 Db 2999 CGACTTTCAATTTCTACTCTCTCTCCCTCGCAAGGCTTCATCGTACTTTGACTGGAAACAC 3058
 QY 383 rThrIleAlaCysTyrValAlaIleLeuIleGlyLeuCysPheThrLeuValLeuAl 403
 Db 3059 GACTATCGTGTGTATGTGGCCATTTCTATCGGTCTCTGCTTCACTCTCTGTCTGCTGCG 3118
 QY 403 aValPheLysArgAlaLeuProAlaLeu -GlnPheProPheSerProAspSerPhePheT 423
 Db 3119 CGTCTTCAACAGAGCACTCCCGCTCTGCCAATTTTCCATTTTCTCGGACTCAATTTT 3178
 QY 423 hrPheValProAlaGlySerSerProHisLeuLeuHisLysSerLeuLysSerValTyrT 443
 Db 3179 CTTTTGTACCCGCTGTGATCATCACCCCATTTGTTACACAAGTCTCTCAAAAAGTGTATT 3238

QY 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh 451
 Db 1617 ATGGACCAATTAGCATTCCTCAATTTATATCTAGCATATTTCGGTTAGATCCCATG 1676
 QY 451 eLeuCysIleIle 455
 Db 1677 GATGTTCTCTT 1689

RESULT 3

US-08-706-344C-27
 ; Sequence 27, Application US/08706344C
 ; Patent No. 6248555
 ; GENERAL INFORMATION:
 ; APPLICANT: TANZI, RUDOLPH
 ; APPLICANT: WASCO, WILMA
 ; TITLE OF INVENTION: Genetic Alterations Related To Familial
 ; TITLE OF INVENTION: Alzheimer's Disease
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVENUE, SUITE 600
 ; CITY: WASHINGTON
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005-3934
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/706,344C
 ; FILING DATE: 30-AUG-1996
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/003,054
 ; FILING DATE: 31-AUG-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KIM, JUDITH U.
 ; REGISTRATION NUMBER: 40,679
 ; REFERENCE/DOCKET NUMBER: 0609.4180001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-2600
 ; TELEFAX: 202-371-2540
 ; INFORMATION FOR SEQ ID NO: 27:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2765 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 249..1649
 ; US-08-706-344C-27

Alignment Scores:
 Pred. No.: 2,318-104 Length: 2765
 Score: 994.50 Matches: 215
 Percent Similarity: 65.17% Conservative: 75
 Best Local Similarity: 48.31% Mismatches: 107
 Query Match: 41.82% Indels: 48
 DB: 3 Gaps: 8

US-09-043-944-1 (1-461) x US-08-706-344C-27 (1-2765)

QY 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValrle 51
 Db 438 GAGCAAGATGAGCAAGAGATGAGGAGCTGACATTGAAATATGGCCCAAGCATGTGATC 497
 QY 52 HisLeuPheValProValSerLeuCysMetAlaLeuValPheThrMetAsnThrle 71
 Db 1500 TTGTGCTTACATTATTACTCTCTCCATTTTCAAGAAAGCATTCGCCAGCTCTTCCAATC 1559

Db 498 ATGCTCTTTTGCCTCGTACTCTCTGCATGGTGGTGGTCTGCTGCTACCATTAAGTCAGTC 557
 QY 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
 Db 558 AGCTTTTATACCGGAGGATGGGAG--CTAATCTATACCCCATTCACAGAAATAC 614
 QY 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
 Db 615 GAGACTGTGGCCAGAGAGCCCTGCACCTCAATCTGATGCTGCCATCATGATCAGTGTC 674
 QY 112 ValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuIle 131
 Db 675 ATTGTTGTGATGACTATCTCTCTGGTGTCTGTATAAATACAGGTGCTATAAGGTGATC 734
 QY 132 HisGlyTrpLeuIleValSerSerPheLeuLeuPheLeuPheThrThrIleTyrVal 151
 Db 735 CATGCTTGGCTTATATATCACTCTATTGTGTGCTGTTCTTTTTCATCTCAITACTTG 794
 QY 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
 Db 795 GGGGAAGTGTATAAAACCTATAACGTTGCTGTGACTACATTACTTGTGCACCTCCTGATC 854
 QY 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTrpLysGlyProLeuArgLeuGln 191
 Db 855 TGAATTTTGGTGTGGGAATGATTTCCATCTACTGGAAGGTCCACTTGCAGCTCCAG 914
 QY 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
 Db 915 CAGGCATATCTCATTTATGATGATGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCT 974
 QY 212 GluTrpThrValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeu 231
 Db 975 GAATGGACTGCGTGGCTCATCTTGGCTGTGATTTAGTATATGATTTAGTGGCTGTTG 1034
 QY 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
 Db 1035 CGTCCGAAGTCCACTTCGTATGCTGTTGTAACAGCTCAGAGAGAAATGAAAGCCTT 1094
 QY 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
 Db 1095 TTTCCAGCTCTCATTTTACTCTCTCAACATGTTGCTG--TTGGTGAATATGGCA 1145
 QY 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
 Db 1146 GAAGAGACCCGGAAGCTCAAGGAGAGATTCAAAATTC----- 1187
 QY 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
 Db 1188 -----AAGCATATATGCA 1199
 QY 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331
 Db 1200 GAAAGCACAGAAAGGAGTCAACAGACTGTTGAGAGAAATGATGATGGCGGTTCAAT 1259
 QY 332 ValGlu-----ArgGluLeuAlaAlaGlu 339
 Db 1260 GAGGAATGGGAAGCCAGAGGAGACTCATCTAGGCGCTCATCGCTCTACCTGAGTCA 1319
 QY 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
 Db 1320 CGAGCTGCTGTCCAGGAACCTTTCCAGCAGTATCTCTCGTGGTGAAGACCCAGAGAAAG 1379
 QY 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
 Db 1380 GGAGTAAACTGGATTTGGAGATTTCTATTCTACAGTGTCTGTTGGTAAAGCTCA 1439
 QY 377 SerTyrPhe-----AspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIle 394
 Db 1440 GCAACAGCCAGTGGAGACTGGAACAACAACATAGCTGTTTCGTAGCCATAATTAATGCT 1499
 QY 395 LeuCysPheThrLeuValLeuLeuAlaValPheLysArgAlaLeuProAlaLeuGlnPh 414
 Db 1500 TTGTGCTTACATTATTACTCTCTCCATTTTCAAGAAAGCATTCGCCAGCTCTTCCAATC 1559

Db 1560 TCCA--TCACCTTTGGGCTTTTCTTCTACTTTTGGCCACAGATTATCTTGTACAGCTTTT 1616
 Qy 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh 451
 Db 1617 ATGGACCAATTAGCATTCACATTTATATCTAGCATATTTGGGTTAGAAATCCCATG 1676
 Qy 451 eleuCysIlelle 455
 Db 1677 GATGTTTCTTCTT 1689

RESULT 9

US-08-592-541-133
 ; Sequence 133, Application US/08592541
 ; Patent No. 5986054
 ; GENERAL INFORMATION:
 ; APPLICANT: ST. GEORGE-HYSLOP, PETER H
 ; APPLICANT: ROMMENS, JOHANNA M
 ; APPLICANT: FRASER, PAUL E
 ; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
 ; TO ALZHEIMER'S DISEASE
 ; NUMBER OF SEQUENCES: 183
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: TESTA, HURWITZ & THIBEAULT
 ; STREET: High Street Tower - 125 High Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/592,541
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pitcher, Edmund R.
 ; TELEPHONE: (617) 248-7000
 ; TELEFAX: (617) 248-7100
 ; INFORMATION FOR SEQ ID NO: 133:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2791 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-592-541-133

Alignment Scores:
 Pred. No.: 3,06e-104 Length: 2791
 Score: 993.50 Matches: 216
 Percent Similarity: 65.84% Conservative: 77
 Best Local Similarity: 48.54% Mismatches: 104
 Query Match: 41.78% Indels: 48
 DB: 2 Gaps: 9

US-09-043-944-1 (1-461) x US-08-592-541-133 (1-2791)

Qy 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValIle 51
 Db 438 GAGCAGATGAGGAGAGAGATGAGGAGCTGACATTGAAATATGGCGCCAGCATGTGATC 497
 Qy 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrIle 71
 Db 498 ATGCTTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTGGCTTACCATTAAGTCAGTC 557
 Qy 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
 Db 558 AGCTTTTATACCGGAGGATGGGAG---CTAATCTATATACCCCATTCACAGAGATACC 614

Qy 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
 Db 615 GAGACTGTGGCCAGAGAGCCCTGCACTCAATCTGAATGCTGCCATCATGACATGTC 674
 Qy 112 ValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuIle 131
 Db 675 ATTGTTGTCATGACTATCTCTCTGGTGTCTGTATATAATACAGGTGCTATAAGGTGTCATC 734
 Qy 132 HisGlyTripleIleValSerSerPheLeuLeuLeuPheLeuPheThrIleTyrVal 151
 Db 735 CATGCCCTGGCTTATATATCATCTTATTGTTGCTGTTCTTTTTCATTATTACTTTG 794
 Qy 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
 Db 795 GGGGAAGTGTTTAAACCTATAACGTTGCTGTGGACTACATTACTGTTCACCTCCTGATC 854
 Qy 172 GlyAsnTyrGlyValLeuGlyMetCysIleHisTrpLysGlyProLeuArgLeuGln 191
 Db 855 TGGAAATTTTGTGTGTGGGAATGATTTCATTCCTGAAAGGTCCACTTCGACTCCAG 914
 Qy 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
 Db 915 CAGGCATATCTCATTTATGATTAGTAGCCCTCATGGCCCTGGTGTATTATCAAGTACCTCCCT 974
 Qy 212 GluTrpThrValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeu 231
 Db 975 GAATGGACTGCTGGTGCATCTTGGCTGTGATTTCAGTATATGATTAGTGGCTGTTTG 1034
 Qy 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
 Db 1035 TGTCCGAAGTCCACTTCTGTATGTGTGTTGAAACAGCTCAGAGAGAATGAAGCGCTT 1094
 Qy 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
 Db 1095 TTTCCAGCTCTCATTTACTCTCAACATGGTGTGG-----TTGGTGAATATGGCA 1145
 Qy 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
 Db 1146 GAAGGAGACCCGGAAGCTCAAGGAGAGATATCCAAAAATTC-----1187
 Qy 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
 Db 1188 -----AAGTATAATGCAGAAAGCACAGAAAG-----1214
 Qy 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331
 Db 1215 -----GAGTCACAAAGACACTGTTGCAGAGAATGATGCGGGTTTCAGT 1259
 Qy 332 ValGlu-----ArgGluLeuAlaAlaGlu 339
 Db 1260 GAGGAATGGGAAGCCAGAGGGACAGTCATCTAGGGCCCTCATCGCTCTACACCTGAGTCA 1319
 Qy 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
 Db 1320 CGAGCTGCTGCCAGGAACCTTTCAGCAGCATCTCTCGCTGGTGAAGACCCAGAGAAAG 1379
 Qy 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
 Db 1380 GGAGTAAACTTGGATTGGAGATTTCATTTCTACRGTTGTTCTGGTGAAGGCTCA 1439
 Qy 377 SerTyrPhe-----AspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
 Db 1440 GCAACAGCCAGTGGAGACTGGAAACAACCATAGCTGTTTCGTAGCCATATTAATTGTT 1499
 Qy 395 LeuCysPheThrLeuValLeuLeuAlaValPheLysArgAlaLeuProAlaLeuGlnPh 414
 Db 1500 TTGTGCTTACATTTATTACTCTTCCTTTCATTTTCAAGAAAGCATTCGACCTCTCCAATC 1559
 Qy 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeu 434
 Db 1560 TCCA---TACCTTTGGGCTGTTTCTTACTTTCACACAGATTAICTTGTAGCGCTTTT 1616
 Qy 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPh 451


```

Db 1617 ATGGACCAATAGCATTCACCAATTTTATATCTAGCATATTTGCGGTTAGAAATCCCATG 1676
Qy 451 eleuCysleile 455
Db 1677 GATGTTTCTCTT 1689

RESULT 11
US-09-127-480-133
; Sequence 133: Application US/09127480
; Patent No. 6194153
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/127,480
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 133:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2791 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-127-480-133

Alignment Scores:
Pred. No.: 3,06e-104 Length: 2791
Score: 993.50 Matches: 216
Percent Similarity: 65.84% Conservative: 77
Best Local Similarity: 48.54% Mismatches: 104
Query Match: 41.78% Indels: 48
DB: 3 Gaps: 9

US-09-043-944-1 (1-461) x US-09-127-480-133 (1-2791)
Qy 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyValAsnHisValile 51
Db 438 GAGCAAGATGAGGAAGAGATGAGGAGCTGACATTTGAAATATGCGCCCAAGCATGTGATC 497
Qy 52 HisLeuPheValProValSerLeuGluGluAlaLeuValValPheThrMetAsnThrIle 71
Db 498 ATGCTCTTGTCCCTGACTCTGTCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 557
Qy 72 ThrPheTyrSerGlnAsnAsnGlyArgHisLeuLeuLeuSerHisProPheValArgGluThr 91
Db 558 ACCTTTTATACCGGAGAGATGGGAG---CTAATCTATACCCCATTCACAGAGATACC 614

```

```

Qy 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
Db 615 GAGACTGTGGCCAGAGAGCCCTGCACTCAATCTGAATGCTGCCATCATGATCAGTGTC 674
Qy 112 ValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuIle 131
Db 675 ATTGTTGTGATGACTATCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 734
Qy 132 HisGlyTyrLeuIleValSerPheLeuLeuLeuPheLeuPheThrIleTyrVal 151
Db 735 CATGCTGTGCTTATATATCACTCTATTTGTTGCTGTTCTTTTTCATTCATTCATTCATTC 794
Qy 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuValLeuPheGlyLeu 171
Db 795 GGGGAAAGTGTAAACCTATAACGTTGCTGGACTACATTAATCTGTCACCTCTGATC 854
Qy 172 GlyAsnTyrGlyValLeuGlyMetCysIleHisTyrPheGlyProLeuArgGln 191
Db 855 TGGAAATTTTGGTGTGGTGGGAATGATTTCCATTCCTGGAAGGTCCTTCGACTCCAG 914
Qy 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrValPro 211
Db 915 CAGGCATATCTCATTAATGATAGTGCCTCATGSCCTGCTGTTTATCAAGTACCTCCCT 974
Qy 212 GluTyrThrValTrpPheValLeuPheValIleSerValTrpAspLeuValAlaValLeu 231
Db 975 GAATGAGCTGCTGCTGCTCATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1034
Qy 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
Db 1035 TGTCGGAAGGTCCTTCTGATGCTGTTGAAACAGCTCAGGAGAGAAATGAAAGCTT 1094
Qy 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
Db 1095 TTTCCAGCTCTCATTTACTCTCAACAATGGTGTGG-----TTGGTGAATATGGCA 1145
Qy 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
Db 1146 GAAGGAGACCCGGAAGCTCAAGAGGAGATATCAAAAATTC----- 1187
Qy 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
Db 1188 -----AAGTATAATGCAGAAAGCACAGAAAG----- 1214
Qy 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331
Db 1215 -----GAGTCACAAGACACTGTTGTCAGAGAAATGATGCGGGTTTCAGT 1259
Qy 332 ValGlu-----ArgGluLeuAlaAlaGlu 339
Db 1260 GAGCAATGGGAAGCCAGAGGACAGTCACTAGGCGCTCATCGCTCTACCTGAGTCA 1319
Qy 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
Db 1320 CGAGCTGCTGCCAGGAATTTTCCAGCAGTATCTCTGCTGGTGAAGACCCAGAGGAAGG 1379
Qy 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
Db 1380 GGAGTAAACTTGGATTGGAGATTTCATTTCTACAGTGTCTGTTGTTGTTGTTGTTGTTGTT 1439
Qy 377 SerTyrPhe-----AspTrpAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
Db 1440 GCAACAGCCAGTGGAGACTGGAAACAACCATAGCTGTTTCGTAGCCATATTAATGTGT 1499
Qy 395 LeuCysPheThrLeuValLeuLeuAlaValPheLysArgAlaLeuProAlaLeuGlnPhe 414
Db 1500 TTGTGCTTACATTAATTTACTCTTGCCATTTTCAAGAAAGCATTCGACCTCTTCCAATC 1559
Qy 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLeu 434
Db 1560 TCCA---TCACCTTTGGGCTGTTTCTACTTGTCCACAGATTAATCTGTACAGCTTTT 1616
Qy 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe-----LeuProPhe 451

```


QY 434 uHisLysSerLeuLysSerValTyrTrpIleAsnSerLeuPhe-----LeuProPh 451
Db 1617 ATGACCAATTAGCAATTCATCAATTTATATCTAGCAATTTTCGGTTAGAAATCCCATG 1676

QY 451 eLeuCysIlelle 455
Db 1677 GATGTTTCTCTT 1689

RESULT 14
US-08-496-841C-133
; Sequence 133, Application US/08496841C
; Patent No. 6210919
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; ROMMENS, JOHANNA M
; FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby, PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/496,841C
; FILING DATE: 28-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul F. Fehner, Ph.D.
; REGISTRATION NUMBER: 35,135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; INFORMATION FOR SEQ ID NO: 133:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 133:
US-08-496-841C-133

Alignment Scores:
Pred. No.: 3,06e-104 Length: 2792
Score: 993.50 Matches: 216
Percent Similarity: 65.84% Conservative: 77
Best Local Similarity: 48.54% Mismatches: 104
Query Match: 41.78% Indels: 48
Gaps: 9

US-09-043-944-1 (1-461) x US-08-496-841C-133 (1-2792)

QY 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValile 51
Db 438 GAGCAAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGCGCGCAAGCATGTGATC 497

QY 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrile 71
Db 498 ATGCTCTTTGTCCTGTGACTCTCTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557

QY 72 ThrPheTyrSerGlnAsnAspGlyArgHisLeuLeuSerHisProPheValArgGluThr 91
Db 558 AGCTTTTATACCCGGAAGGATGGGAG---CTAATCTATACCCCATTTACAGAGAATACC 614

QY 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
Db 615 GAGACTGTGGCCAGAGAGCCCTGCTCACTCAATCTCTGAATGCTGCCATCATGATCAGTGC 674

QY 112 ValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeu 131
Db 675 ATTGTGTGATGACTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734

QY 132 HisGlyTyrLeuIleValSerSerPheLeuLeuLeuPheLeuPheThrThrThrThrVal 151
Db 735 CATGCTGCTTATTATATCATCTCTATGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 794

QY 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
Db 795 GGGGAAGTGTGTTAAACCTATAACCTGCTGGGACTACATTACTGCTGCTGCTGCTGCTGCTGCT 854

QY 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTyrLysGlyProLeuArgGln 191
Db 855 TGGAAATTTTGGTGTGGTGGGAATGATTTCCATCTACTGGAAAGGTCCTCTGCTGCTGCTGCT 914

QY 192 GlnPheTyrLeuIleThrMetSerAlaLeuMetAlaLeuValPheIleLysTyrLeuPro 211
Db 915 CAGGCATATCTCATTTATGATTAGTCCCTCATGCGCCCTGGTGTGTTATCAAGTACCTCCCT 974

QY 212 GluTyrThrValTyrPheValLeuPheValIleSerValTyrPheValLeuValLeu 231
Db 975 GAATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034

QY 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluPro 251
Db 1035 TGTCCGAAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1094

QY 252 PheProAlaLeuIleTyrSerSerGlyValIleTyrProTyrValLeuThrAlaVal 271
Db 1095 TTTCCAGCTCTCATTTACTCTCTCAACATGGGTGG---TTGCTGAATATGGCA 1145

QY 272 GluAsnThrThrAspProGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
Db 1146 GAAGGAGACCCGGAAGCTCAAGGAGAGTATCCAAAAATTC--- 1187

QY 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
Db 1188 -----AAGTATAATGCAGAAAGCAGAAAGG----- 1214

QY 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGluAsnSerGlyValArg 331
Db 1215 -----GAGTCACAGACACTGTTGCAGAGAATGATGATGGGGGTTCCAGT 1259

QY 332 ValGlu----- 339
Db 1260 GAGGAATGGGAAGCCCGAGGAGCAGTCACTAGGGCCCTCATCGCTCTACCTCAGTCA 1319

QY 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
Db 1320 CGAGCTGCTGCTCCAGGAACCTTTCCAGCAGATATCTCGCTGGTGAAGACCCAGAGAAAGG 1379

QY 357 GlyValLysLeuGlyLysPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
Db 1380 GGAGTAAACCTTGGATTGGGAGATTTCATTTCTACAGTGTCTGTTGGTGAAGACCTCA 1439

QY 377 SerTyrPhe-----AspTyrAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
Db 1440 GCAACAGCCAGTGGAGACTGGGAACAACCATAGCTCTGTTTCGTAGCCATATTAATGTT 1499

QY 395 LeuCysPheThrLeuValLeuLeuAlaValPheLysArgAlaLeuProAlaLeu-GlnPh 414
Db 1500 TTGTGCTCTTACATTTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1559

QY 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeu 434
Db 1560 TCCA---TCACCTTTGGGCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1616

QY 434 uHisLysSerLeuLysSerValTyrTrpIleAsnSerLeuPhe-----LeuProPh 451


```
Db 1617 ATGACCAATAGCATTCATCAATTTATATATAGCATATTTCGGTTAGAAATCCCATG 1676
QY 451 eLeuCysIleile 455
Db 1677 GATGTTCTTCCTT 1689

RESULT 15
US-08-670-964-3
; Sequence 3, Application US/08670964
; Patent No. 6010874
; GENERAL INFORMATION:
; APPLICANT: Hardy, John A.
; TITLE OF INVENTION: EARLY ONSET ALZHEIMER'S DISEASE
; TITLE OF INVENTION: GENE AND GENE PRODUCTS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - UW2220; P.O. Box 15
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,964
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,142
; FILING DATE: 13-JUL-1995
; APPLICATION NUMBER: 60/001,501
; FILING DATE: 18-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: P50358
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1750 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-670-964-3

Alignment Scores:
Pred. No.: 2,68e-104 Length: 1750
Score: 991.00 Matches: 215
Percent Similarity: 65.98% Conservative: 72
Best Local Similarity: 49.43% Mismatches: 103
Query Match: 41.67% Indels: 45
DB: 3 Gaps: 8

US-09-043-944-1 (1-461) x US-08-670-964-3 (1-1750)
QY 32 GlnGluAspGluAsnValValGluGluAlaGluLeuLysTyrGlyAlaSerHisValile 51
Db 351 GAGCAAGATGAGGAGAGATGAGGAGCTGACATTGAAATATGGCCCAAGCATGTGATC 410
QY 52 HisLeuPheValProValSerLeuCysMetAlaLeuValValPheThrMetAsnThrIle 71
Db 411 ATGCTCTTTGCTGACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 470
QY 72 ThrPheTyrSerGlnAsnAsnGlyAArgHisLeuLeuSerHisProPheValArgGluThr 91
```

```
Db 471 AGCTTTTATACCCGGAAGGATGGGCAG---CTAATCTATATACCCCATTCACAGAGAATACC 527
QY 92 AspSerIleValGluLysGlyLeuMetSerLeuGlyAsnAlaLeuValMetLeuCysVal 111
Db 528 GAGACTGTGGCCAGAGAGCCCTGCACTCAATCTGTAATGCTGCCATCATGATCAGTGTG 587
QY 112 ValValLeuMetThrValLeuLeuIleValPheTyrLysTyrLysPheTyrLysLeuLeu 131
Db 588 ATTGTTGTCATGACTATCTCTCTCTGTTCTGTATAAATACAGGTCTATAAGGTATC 647
QY 132 HisGlyTyrLeuLeuValSerPheLeuLeuLeuPheLeuPheThrThrIleTyrVal 151
Db 648 CATGCTGCTGCTTATATATCATCTCTATTGTTGCTGCTCTCTCTCTCTCTCTCTCT 707
QY 152 GlnGluValLeuLysSerPheAspValSerProSerAlaLeuLeuValLeuPheGlyLeu 171
Db 708 GGGGAAGTGTATAAACCTATAAGTGTGCTGTGAGTACATTAATGTTGCTACTCTGATC 767
QY 172 GlyAsnTyrGlyValLeuGlyMetMetCysIleHisTyrLysGlyProLeuArgLeuGln 191
Db 768 TGGAAATTTGGTGTGGTGGGAATGATTTCCATTCACCTGGAAAGGTCCACTTCGACTCC 827
QY 192 GlnPheTyrLeuLeuThrMetSerAlaLeuMetAlaLeuValPheLeuLysTyrLeuPro 211
Db 828 CAGGCATATCTCATTTATGATTAGTGCCTCATCGCCTGCTGTTTATCAAGTACCTCCCT 887
QY 212 GluTyrThrValTyrPheValLeuPheValIleSerValTyrPheValAlaValLeu 231
Db 888 GAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 947
QY 232 ThrProLysGlyProLeuArgTyrLeuValGluThrAlaGlnGluArgAsnGluProIle 251
Db 948 TGTCCGAAAGGTCCACTTCGTATGCTGTTGTTGAAACAGCTCAGGAGAGAAATGAACGCT 1007
QY 252 PheProAlaLeuLeuTyrSerSerGlyValIleTyrProTyrValLeuValThrAlaVal 271
Db 1008 TTTCCAGCTCTCATTTACTCTCAACATGTTGTGG-----TTGTGAAATATGGCA 1058
QY 272 GluAsnThrThrAspProArgGluProThrSerSerAspSerAsnThrSerThrAlaPhe 291
Db 1059 GAAGGAGAGCCCGAAGCTCAAGGAGAGTATCAAAAATTC----- 1100
QY 292 ProGlyGluAlaSerCysSerSerGluThrProLysArgProLysValLysArgIlePro 311
Db 1101 -----AAGTATAATGCAAGAACACAGCAAGG----- 1127
QY 312 GlnLysValGlnIleGluSerAsnThrThrAlaSerThrThrGlnAsnSerGlyValArg 331
Db 1128 -----GAGTCACAAGACACTGTTGTCAGAGAATGATGCGCGGTTCAGT 1172
QY 332 ValGlu-----ArgGluLeuAlaAlaGlu 339
Db 1173 GAGGAATGGGAAGCCAGAGGAGACAGTCATCTAGGGCCCTCATCGCTCTACACCTGAGTCA 1232
QY 340 ArgProThrValGlnAspAlaAsn-----PheHisArgHisGluGluGluArg 356
Db 1233 CGAGCTGCTGTCCAGGAACCTTCCAGCAGTATCTCGCTGGTGAAGACCCAGAGGAAAGG 1292
QY 357 GlyValLysLeuGlyLeuGlyAspPheIlePheTyrSerValLeuLeuGlyLysAlaSer 376
Db 1293 GGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTCTGTTGTTGTTGTTGTTGTT 1352
QY 377 SerTyrPhe-----AspTyrAsnThrThrIleAlaCysTyrValAlaIleLeuIleGly 394
Db 1353 GCAACAGCCAGTGGGAGACTGGACACCAACCATAGCTGTTTGTAGCCCATTAATTTGTT 1412
QY 395 LeuCysPheThrLeuValLeuLeuAlaValPheLysArgAlaLeuProAlaLeuGlnPhe 414
Db 1413 TTGTGCTTACATTTACTCTTGCCATTTTCAAGAAAGCATTTGCCAGCTCTTCCAAATC 1472
QY 414 eProPheSerProAspSerPhePheThrPheValProAlaGlySerSerProHisLeuLeu 434
```

Db 1473 TCCA---TCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGACAGCCTTTT 1529
 Qy 434 uHisLysSerLeuLysSerValTyrTyrIleAsnSerLeuPhe 448
 Db 1530 ATGGACCAATTAGCATTCCATCAATTTTATATATCTAGCATATTT 1572

Search completed: December 11, 2003, 07:43:07
 Job time : 122 secs